



# Methodological Challenges to the Evaluation of Interventions for Foster/Kinship Carers and Children: A Systematic Review

Adam Dickes<sup>1</sup> · Jacqueline Kemmis-Riggs<sup>1</sup> · John McAloon<sup>1</sup>

Published online: 24 November 2017  
© Springer Science+Business Media, LLC, part of Springer Nature 2017

## Abstract

Caregivers of children in alternative care face a complex range of challenges that can result in placement disruption and poor long-term outcomes. Interventions designed to help carers meet these challenges report positive outcomes. Nevertheless, several reviewers have reported these positive results may be mitigated by limitations in trial methodology. This review aims to systematically review these methodological challenges and limitations, to provide an analysis of the current state of the evidence base for these interventions. A systematic review was conducted into the methods used to evaluate the effectiveness of psychosocial interventions for foster and kinship carers. Limitations relating to internal validity, external validity and clinical heterogeneity were identified and synthesised. Seventeen studies met inclusion criteria. The quality of methods used in the included studies is mixed, with high and unknown levels of bias in the majority of trials. Heterogeneity in participant characteristics, intervention aims and outcome measures across interventions reflect the diversity of carer and child needs and make it difficult to generalise results or compare and synthesise the efficacy of different interventions. These factors limit the application of trial results to evidence-based clinical practice. The diverse and complex needs of this population present significant challenges to robustly evaluating interventions for foster/kinship families. Participant needs, theoretical approaches, intervention aims and outcome measures need to be better coordinated, both within trials and across the field. Exploratory research should be used to generate focussed and concrete hypotheses that can be robustly tested in high-quality randomised controlled trials. Protocol registration number: CRD 42017048415.

**Keywords** Foster care · Alternative care · Looked after children · Maltreatment · Evidence-based practice · Methodology

## Introduction

Psychosocial interventions for foster/kinship families are implemented, disseminated and evaluated in an exceptionally complex and challenging context. It is a context that includes the wide-ranging and heterogeneous difficulties of children in alternative care, the varying capacity of carers to meet these children's needs and the multifaceted child protection system itself. The importance of these factors, which have been well established in research, is essential in understanding the methodological limitations and challenges analysed in this review.

Children's experience of maltreatment prior to their entry into alternative care has repercussions across a wide spectrum of domains essential to their well-being and development. Enduring and complex trauma is associated with increasing complexity of symptoms (Cloitre et al. 2009). Abuse and neglect, occurring across essential development periods, has been associated with dysfunctions in neurobiological systems associated with physiological, arousal, relational and executive brain function (Perry 2009) resulting in poor physical health, hyperarousal, attachment and attentional problems (Ford and Courtois 2013; Marti Haidar 2013; Fisher et al. 2000). As a result, children in care have recognised difficulties with emotion and behavioural regulation (Greenberg et al. 2001; Appleyard et al. 2005), an increased prevalence of mental disorders including depression, anxiety and posttraumatic stress disorder (Leenarts et al. 2013; Ford and Courtois 2013) and delays in cognitive and academic functioning (Jacobsen et al. 2013; Gutman et al. 2003).

---

✉ Adam Dickes  
Adam.P.Dickes@student.uts.edu.au

<sup>1</sup> Centre for Child and Adolescent Mental Health, University of Technology Sydney, Level 4 Building 7, 67 Thomas St, Ultimo, NSW 2007, Australia

When children are in family-based alternative care (i.e. foster or kinship care), the principal responsibility for meeting this wide array of needs falls on children's carers. Ideally, foster/kinship carers provide a caring and consistent environment that helps to remediate the effects of maltreatment, with a wide range of research indicating that stable, early placement predicts better long-term outcomes across a number of domains (James et al. 2004; Sinclair et al. 2005; Oosterman et al. 2007; Rubin et al. 2007). The provision of care, however, faces significant challenges, including the behavioural problems identified as the strongest predictor of placement breakdown (Oosterman et al. 2007) which is in turn a strong predictor of subsequent placement breakdown. Placement instability is common in the foster care system, with a significant number of foster children moving through several placements in their childhood (Sinclair et al. 2005; James et al. 2004). Unsurprisingly, this continued instability exacerbates problems in developmental, behavioural and mental health domains (Rubin et al. 2007; Ryan and Testa 2005).

In addition to challenges related to their foster child's needs, foster/kinship carers are required to navigate a complex, poorly resourced and at times poorly functioning child welfare system, which some carers report as their greatest single source of stress (Buehler et al. 2003). An integrative review of research on foster carers' experience found foster carers reporting a range of stresses including poor relationships and disagreements with case workers, and a lack of effective resources and supports (Blythe et al. 2014). Navigating these challenges may be further complicated by the conflicted identity foster/kinship carers may experience as fulfilling both professional and parental roles, with concomitant issues around attachment, uncertainty and commitment (Blythe et al. 2014). This range of barriers to the provision of stable and caring foster/kinship placements may, in some part, explain the generally poor long-term outcomes for children in alternative care. Reviewing 30 years of longitudinal data, Goemans et al. (2015) found that adaptive function and internalising and externalising behaviour problems remained unchanged during children's time in foster care, with poorer outcomes associated with longer time spent in care.

### Interventions for Foster/Kinship Carers

Researchers and clinicians have responded to this array of challenges by developing a wide range of interventions for family-based alternative care (i.e. foster/kinship care). Typically, such interventions are designed to increase foster/kinship carers' skills in an effort both to better manage the sequelae of children's maltreatment and disrupted attachment and their capacity to provide stable, responsive and consistent care. Increasingly, these skills and capacities have come to be seen as essential to effective foster/kinship care

(Turner et al. 2007; Everson-Hock et al. 2012). In keeping with the growing demands of evidence-based practice, foster/kinship family interventions have been evaluated using increasingly sophisticated methodologies over recent years, with a growing use of randomised controlled trials (RCTs). Evidence from these trials has been synthesised in a range of reviews, systematic and empirical reviews and meta-analyses. Review authors have reported a range of methodological limitations that challenge their ability to draw robust conclusions and make clear recommendations about the effectiveness of these interventions. Limitations relating to internal validity included mixed study quality (Leve et al. 2012; Kerr and Cossar 2014), insufficient sample sizes (Festinger and Baker 2013; Craven and Lee 2006) and lack of replication (Kinsey and Schlösser 2013). Problems relating to outcome measures included the use of non-validated measures (Craven and Lee 2006) and the heterogeneity of outcome measures that made it difficult to adequately synthesise or compare trial results (Leve et al. 2012). Challenges to external validity included trial participants that were not representative of the alternative care population, outcome measures not representative of foster child well-being (Leve et al. 2012; Everson-Hock et al. 2012; Dorsey et al. 2008), limited long-term follow-up (Luke et al. 2014; Tarren-Sweeney 2014), a lack of fidelity assurance (Festinger and Baker 2013) and absence of metrics conveying meaningful clinical change (Tarren-Sweeney 2014). Reviewers also noted significant heterogeneity of clinical factors, including diversity in intervention aims, duration and setting and participant characteristics (Racusin et al. 2005; Everson-Hock et al. 2012; Dorsey et al. 2008; Kerr and Cossar 2014; Tarren-Sweeney 2014), all of which make intervention planning and evaluation more challenging. Two reviewers recommended that future evaluation research needs to pay more attention to the clinical, experiential and cultural context of participants, in order to provide essential information about how these factors lead to differential responses to treatment, including possible harms (Luke et al. 2014; Tarren-Sweeney 2014). One reviewer noted that the limitations in current evaluation research reflect, to some extent, the limitations in our current theoretical frameworks, which do not adequately capture the complex developmental challenges of children in alternative care (Tarren-Sweeney 2014).

Potentially, these methodological limitations represent a significant barrier to evidence-based practice in the field. The function of RCTs is to establish an unbiased measure of an intervention's effectiveness in a specified setting for a specified population (Ioannidis et al. 2008). Fulfilling this objective depends upon internal validity, such that methods and procedures are followed to minimise the risk of bias and limit non-systematic effects on trial outcomes (Clarke and Oxman 2000). In addition, RCTs depend upon external validity, such that the results of trials are generalisable to

a defined group of patients in a specified setting (Rothwell 2005). Thus, constraints to internal validity limit the robustness of clinical trial results, while constraints to external validity limit the generalisability of trials results to clinical practice.

Systematic reviews and meta-analyses of evaluation trials function to systematically synthesise and compare trial results in order to guide clinical practice and policy. Reviewers therefore depend on both the internal and external validity of included trials, and the use of valid criteria to include and exclude trials from review (Ioannidis 2005). In addition, the meaningful comparisons and syntheses of individual trial results depend on a sufficient level of uniformity between trials (Higgins et al. 2011). This is the case because clinical heterogeneity, including differences in participant characteristics, trial methods and measured outcomes, can moderate the magnitude of the intervention effect (West et al. 2010), risking the introduction of non-systematic bias in reviewers' conclusions and, by extension, threatening the internal validity of those conclusions. Finally, because these factors also determine which participants benefit from which interventions in which settings, heterogeneity threatens external validity that allows research conclusions to inform clinical practice and policy.

## Aims of Review

While existing reviews have reported a great number of limitations in the research evaluating interventions for foster/kinship families, their full impact on the evidence base has yet to be identified in a systematic review. As a result, consensus is limited as to which interventions provide robust evidence of efficacy for which problems and participants. In addition, while reviewers have provided a range of helpful recommendations for future research, these recommendations are not based on a systematic review focussed solely on methodological limitations.

This review aims to systematically identify challenges related to the complexity of the field that currently prevent a robust evaluation of interventions that can be synthesised, compared and generalised to foster/kinship carers and children in care. Specifically, it will provide a systematic review of factors that threaten both the internal and external validity of programme evaluations, and the clinical heterogeneity that exists between programmes. Given the improvement in methodological quality over time (Festinger and Baker 2013), this study will focus on the most recent evaluations conducted using the highest available study quality.

The present review will address the following research questions:

1. What are the threats to external validity of randomised trial evaluations?

2. How much clinical heterogeneity is present between randomised trials evaluations?
3. What are the threats to internal validity in randomised interventions for foster families?

## Method

### Inclusion/Eligibility Criteria

#### Types of Studies

Studies were eligible for this review if they were conducted after 1990, were published in a peer-reviewed journal, had more than 20 participants and reported quasi-random or random allocation of participants to control or experimental groups. Only studies that evaluated interventions with a stated aim of improving child well-being were included.

#### Types of Participants

Interventions for foster and kinship carers (with or without their foster children) caring for children under the age of 18. The focus of this review is on interventions primarily directed to family-based alternative care (i.e. foster or kinship care). For this reason, we excluded interventions targeting residential group care, institutional care, children referred to care by the juvenile justice system and interventions that solely targeted adoptive or biological parents.

#### Types of Interventions

Psychosocial interventions that reported primary aims of improving child well-being, including behavioural, mental health, relational and attachment approaches. Specifically, aims related to reducing child behaviour, improving child mental health, child interpersonal skills, child biomarkers of stress, foster/kinship parent–child relationships, carer well-being, parenting skills and placement stability were included. Interventions with a primary aim of improved academic or school functioning were excluded. Programmes that provided additional support outside of the direct intervention, commonly known as 'wrap-around' services, were excluded, as this was beyond the scope of this study.

#### Types of Outcome Measures

Interventions that used at least one psychometrically validated quantitative outcome measure related to child well-being.

## Types of Comparisons

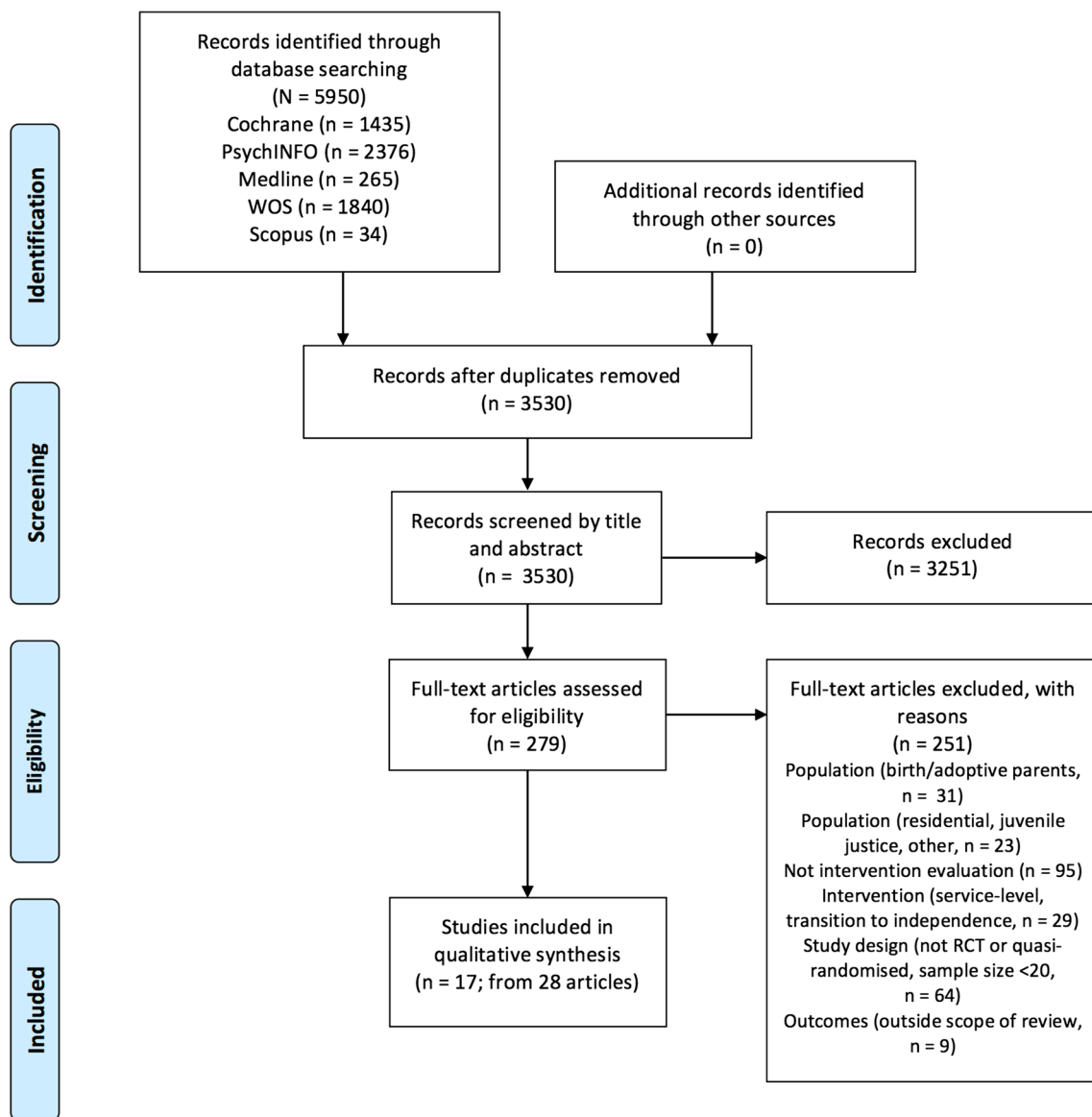
Studies that randomised participants to an intervention group and an active or inactive control group, and reported comparative change in outcomes after the intervention.

## Search Methods for Identification of Studies

The databases CINAHL, Cochrane Library, PsychINFO, Scopus, Medline and Web of Science were searched for articles published between January 1990 and August 2016. Search terms were modified to meet the individual requirements of databases.

## Selection of Studies

Studies were screened and excluded by title by one of the authors (JK). Studies were then screened independently by abstract and by article by two authors (JK and AD). Discrepancies were resolved by consensus after further detailed analysis and reading. The first stage of the project involved a systematic review with qualitative data analysis consistent with the PRISMA-P statement (Moher et al. 2015). Figure 1 provides an overview.



**Fig. 1** Flowchart of selection of studies for inclusion in this review

## Data Extraction and Management

Data were extracted by AD and JK independently. No imputation of missing data was undertaken. Five randomly selected included studies were used to pilot the criteria and extraction process, which were modified after consultation between researchers. Data relating to intervention setting and format, participant characteristics, intervention characteristics and outcomes were extracted from each study. Setting and format data included location, duration, dosage, delivery, format and setting. Data relating to participant characteristics included age, gender, placement history, maltreatment history, foster/kinship carer experience and family characteristics. Data relating to the interventions included theoretical approaches, intervention aims and outcomes measures used to evaluate efficacy. These data categories were chosen due to their importance as predictors of foster child outcomes.

## Risk of Bias

Risk of bias was based on reporting in published journal articles and was evaluated across six domains using the Cochrane risk of bias tool (Higgins et al. 2011). Bias was rated independently by two authors (JK and AD) with inter-rater reliability assessed by an independent researcher (LL). Inter-rater reliability varied across risk of bias domains with good agreement in half of the domains and moderate agreement in the remaining half. The random sequence generation domain had an inter-rater reliability test agreement of 70.6%  $\kappa = 0.512$  ( $p < 0.002$ ). The allocation concealment domain had an inter-rater reliability test agreement of 94.1%  $\kappa = 0.883$  ( $p < 0.001$ ). The performance bias domain had an inter-rater reliability test agreement of 100%  $\kappa = 1$  ( $p < 0.001$ ). The outcome assessment bias received the lowest inter-rater reliability of all domains, with agreement of 70.6%  $\kappa = 0.514$  ( $p < 0.002$ ). The attrition bias domain had an inter-rater reliability test agreement of 88.2%,  $\kappa = 0.79$  ( $p < 0.001$ ). The reporting bias domain had an inter-rater reliability test agreement of 76.5%,  $\kappa = 0.511$  ( $p < 0.001$ ). Reported results are a consensus agreement made where there was some disagreement.

## Results

As shown in Fig. 1, 17 studies were included in the final analysis. Table 1 provides a list of included intervention names and abbreviations. See ‘Appendix’ for further characteristics of the included studies.

**Table 1** Abbreviation for included interventions

Abbreviation	Intervention name	Date
ABC1	Attachment and Bio-behavioral Catch-up	2006
ABC2	Attachment and Bio-behavioral Catch-up	2009
CEBPT	Cognitive Behavioural Parent Training for Externalising Disorders	2012
CBT-PT	Cognitive Behavioural Parent Training Program	2012
FCCT	Foster Carers Communication Training	2016
FFI	Foster Carer–Foster Child Intervention	2016
IY	Incredible Years	2011
IY + CP	Incredible Years Co-parenting Adaptation	2006
KEEP1	Keeping Foster Parents Trained and Supported	1992
KEEP2	Keeping Foster Parents Trained and Supported	2008
KEEP3	Keeping Foster Parents Trained and Supported	2015
KITS	Kids in Transition To School	2012
MSS	Middle School Success	2011
PCIT	Parent–Child Interaction Therapy	2014
PFR	Promoting First Relationships	2012
PMTO	Parent Management Training Oregon	2015
PSB	Promoting Sibling Bonds	2015

## Setting and Format

Eleven of the seventeen studies included in this review were conducted in the USA, three in the UK, two in the Netherlands and one in Romania. All study results were published in peer-reviewed journals. Five were conducted by researchers currently affiliated with the Oregon Social Learning Centre (KEEP1-3, KITS, MSS). All other interventions were evaluated by individual researchers.

Duration, dosage, format and delivery of the interventions varied a great deal. Figure 2 summarises intervention dosage across studies. Duration varied in length from 3 days to 16 weeks. Individual session length varied from 1 to 8 hours, a range of 3 to 52 total sessions and 8–104 total hours. Three interventions were delivered to parent–child dyads in the home (ABC, FFI, PFR), and one was delivered to individual parents in a community setting (PMTO). All remaining interventions utilised group formats in a community setting. Just over half of interventions (53%) were delivered to both foster/kinship carers and children in their care (ABC1, ABC2, FFI, IY + CP, KITS, MSS, PCIT, PFR, PSB) with the remaining interventions delivered to foster/kinship carers only.

## Participant Characteristics

Figure 3 summarises overall levels of reporting of demographic information reported by included studies. All included studies reported some demographic information for children in care and their carers including age, gender,

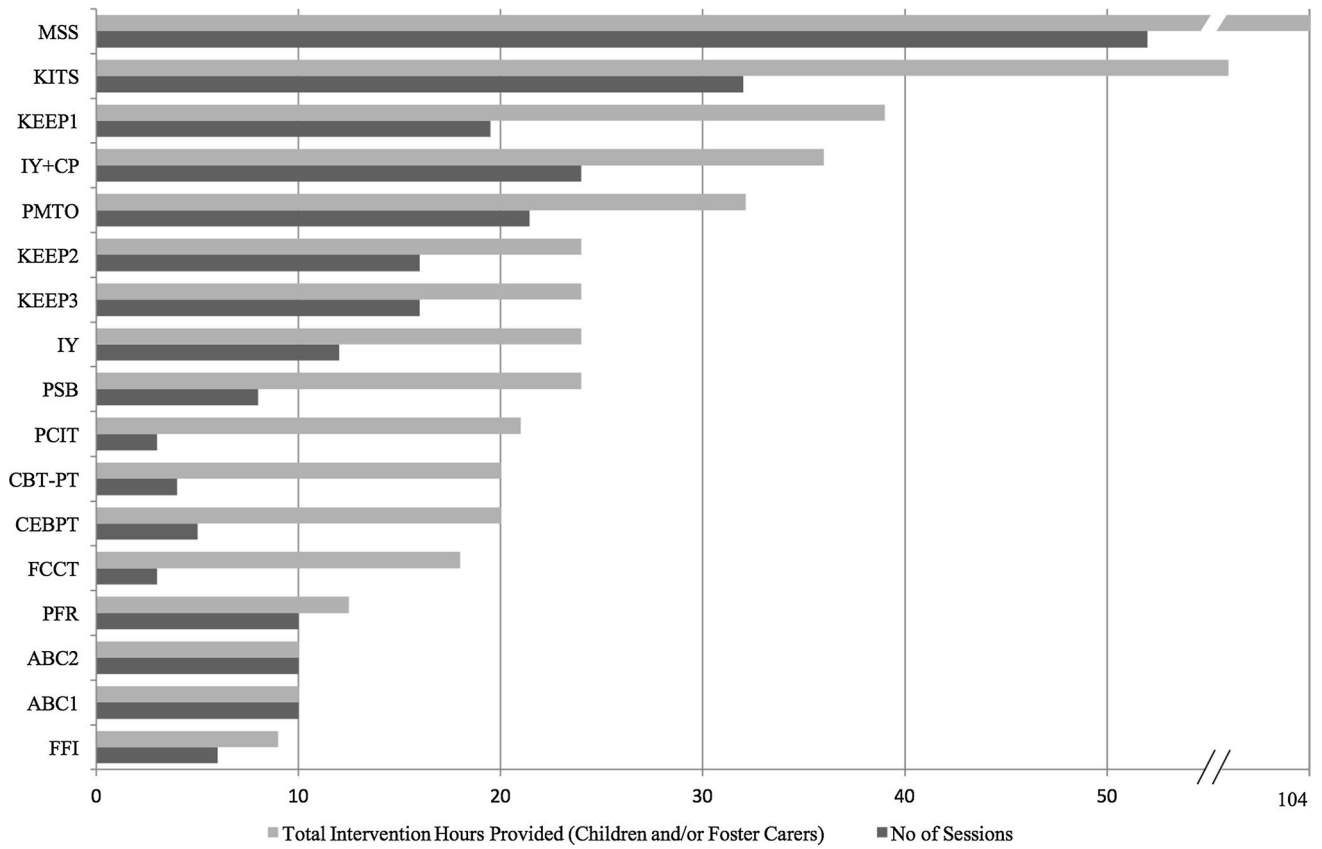


Fig. 2 Duration and length of interventions

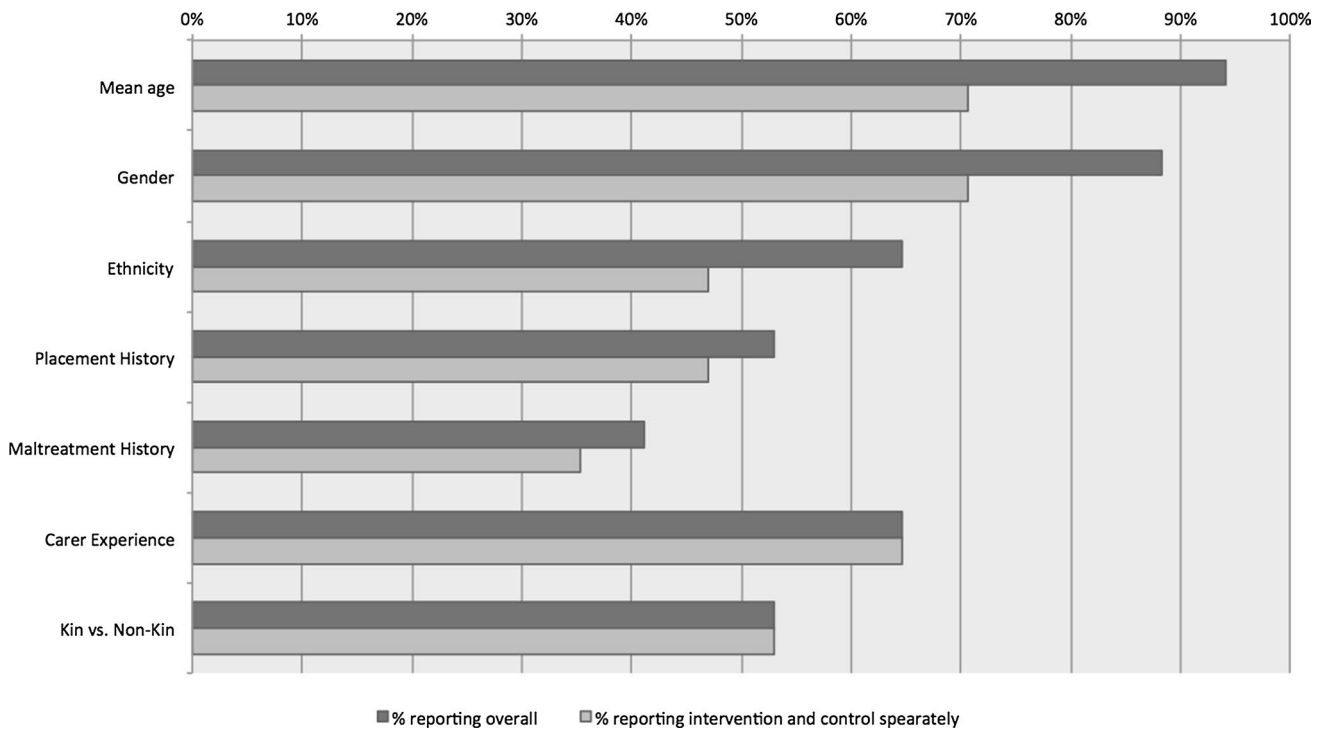


Fig. 3 Reporting of demographic information overall and separately for intervention and control



placement history, maltreatment history, carer experience and family characteristics. Demographic information for foster/kinship carers was generally well reported, while reporting of demographic information for index foster/kinship children was less extensive, despite its identification as an important predictor of child well-being (Redding et al. 2000; Oosterman et al. 2007). Determining the representativeness of participants and degree of heterogeneity in participant characteristics was limited by a lack of complete reporting.

Figure 4 summarises age range and mean age of trial participants. Mean age was the most consistently reported participant characteristic; all but one of the studies (CBT-PT) reported the mean age of the index child at baseline. Mean age of participants varied a great deal across studies, from 18 months to 11 years. Age range of included participants varied from 1 year (ABC1) to 15 years (KEEP1).

Fifteen of seventeen studies (88%) reported on the proportion of male and female index foster/kinship children, with a generally even proportion of male and female participants in most studies. Eleven interventions (65%) reported data relating to ethnicity for children and carers, with significant heterogeneity between studies that ranged from 74%

African American participants (ABC1) to 86% Caucasian (KEEP1).

53% of studies reported data relating to the index child’s placement history, reflecting its recognised importance as a predictor of foster/kinship family functioning and placement stability. Placement history was reported in a wide variety of formats, precluding a systematic analysis or comparison of its role as a mediating variable. Four studies reported the number of children’s prior placements (KEEP1, KEEP2, MSS, PMTO). Placement history was also reported as age at removal (ABC1, PFR, PMTO, MSS), length of current placement (FCCT, IY) and time in alternative care (MSS). Seven studies reported no placement history. Variation in age of removal was evident between interventions aimed at younger children (e.g. ABC) and those including older children (e.g. KEEP2), with studies involving younger children generally demonstrating younger age of removal.

Maltreatment history of the index children was reported in seven interventions (41%). Five studies reported the percentage of children that experienced neglect and abuse (FCCT, KEEP1, MSS, IY – CP, PSB), and one reported the percentage of children with a maltreatment history (FFI). One study (ABC1) reported that all children were removed

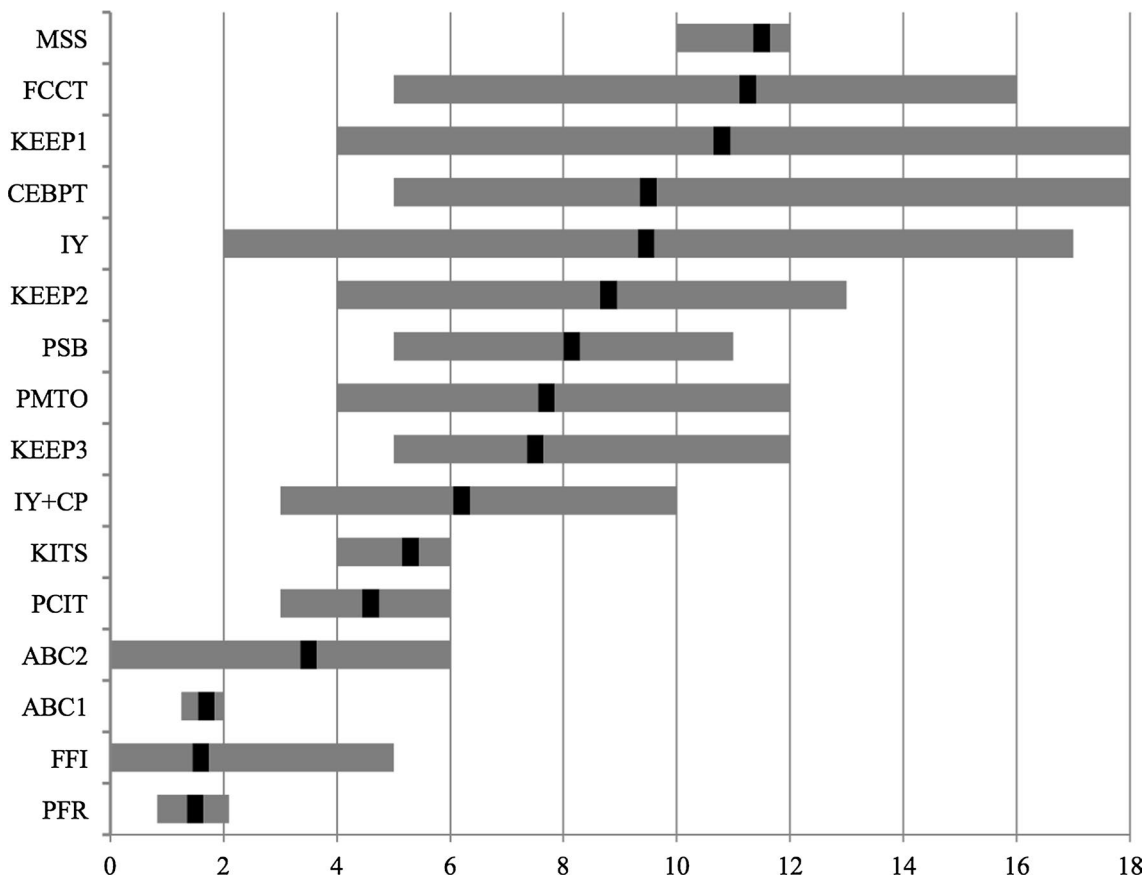


Fig. 4 Mean age and age range of participants in included studies

due either to neglect of children or to psychopathology or incarceration of biological parents. Ten studies reported no information on children's maltreatment history (ABC2, CBT-PT, CEBPT, IY, KEEP2, KEEP3, KITS, PCIT, PFR, PMTO).

Foster/kinship carer characteristics were generally well reported. Years of carer experience was recorded in 13 studies (65%). Eleven studies reported the mean years of carer experience, which ranged from 2 to 8 years (CBT-PT, IY, KEEP1, KEEP2, KEEP3, KITS, MSS, PCIT, PFR, PMTO, PSB). Four studies did not report this variable. Fostering context (kinship or non-kinship care) was recorded in nine studies (53%) and ranged from 0% kin (CEBPT) to 47% kin.

### Theoretical Approaches of Included Studies

Three interventions were described as being based on principles of Attachment Theory (ABC, PFR, FFI). Four were described as being based on the principles of Social Learning Theory (IY, IY + CP, KEEP, PMTO). One intervention was described as being founded on a combination of Attachment and Social Learning Theories (PCIT), one on Cognitive Behavioral principles (CEBPT) and one on a combination of Cognitive Behavioural and Social Learning Theory and (CBT-PT). One intervention was based on a combination of Social learning and Developmental Theory (MSS) and another (PSB) on a combination of Social Learning, Family Systems and Emotion Regulation theories. One intervention (FCCT) did not report the theoretical basis on which the intervention was based.

### Intervention Aims

Reflecting the wide variety of theoretical approaches and the diverse and complex needs of children and their foster/kinship carers, intervention aims exhibited significant variation and complexity, with the majority of interventions stating more than one primary aim. Four interventions were reported as having different primary aims across multiple published studies (ABC1, KEEP2, KITS, MSS).

#### Intervention Aims Related to Improving Carer Capacity and Skills

Several interventions reported a primary aim of improving foster child well-being by improving parenting skills and capacity. Interventions based on attachment principles that involved foster children younger than 6 years stated aims including enhancing caregiver sensitivity and empathy (PFR), helping carers to create an environment that fosters regulatory abilities (ABC1, ABC2), recognising and responding to child stress (FFI) and meeting unmet needs (PFR). Interventions based on Social Learning Theory

involved carers of children up to the age of eighteen and reported aims of reducing behaviour problems and placement disruption by increasing use of positive reinforcement (IY, PMTO, CEBPT, CBT-PT, KEEP2 and KEEP3) on increasing parent involvement (KITS). Three interventions reported aims relating to increasing carers confidence to manage behavioural problems (CBT-PT, IY, IY + CP). One intervention (FCCT) reported an aim to increase foster carers communication skills and confidence. One intervention (PCIT) reported aims of modifying both the responsiveness of the parent, and also developing their behaviour management skills, reflecting its foundation in both Attachment and Social Learning Theory.

#### Intervention Aims Directly Related to Child Behaviour

Reduction in behaviour problems was the most commonly reported intervention aim (KEEP2, CBT-PT, PMTO). Four interventions reported an intervention aim of reducing externalising behaviour (CEBPT, IY + CP, IY, MSS).

#### Intervention Aims Related to Emotion Regulation

One intervention stated emotion regulation as the primary aim of the intervention (ABC1), while another reported the broader domain of self-regulation as one of several aims (MSS).

#### Intervention Aims Related to Placement Outcomes

Four studies reported a reduction in placement disruption, or inversely an increase in placement stability as long-term aims of the intervention (MSS, KEEP1, KEEP2, PFR).

#### Other Intervention Aims

In addition to the aims described above, some interventions had aims specifically related to the target of the intervention. These included development of literacy and pro-social skills (KITS) and the long-term reduction in internalising problems, substance use, delinquency and high-risk sexual behaviour (MSS).

#### Outcomes Measures Used to Evaluate Interventions

The outcomes measures used by included studies are summarised in Table 2. Reported outcomes spanned a broad range of parent and child domains, reflecting the diversity of foster child needs, theoretical approaches and intervention aims. This resulted in a general lack of convergence in outcome measures used: of the 122 outcomes reported across all of the trials, 100 unique measures or scales were used, spanning fifteen domains. On average, studies reported



**Table 2** List of outcomes used to evaluate efficacy in included trials

Domain	Dom. Freq	Outcome measures	Registered	Rep. Freq.																					
				✓	✓	ABC1	ABC2	FCCT	FFI	PCIT	PSB	KEEP1	KEEP2	KEEP3	PMTO	CBT-PT	CEBPT	IY	IY+CP	KITS	MSS				
Placement	9	1. Carer drop out	1																						
		2. Succ. Days	1																						
		3. Changes	1																						
		4. Pos Exits	1																						O
		5. Neg Exits	1																						
		6. Breakdown	1																						
		7. Disruption	1																						
		8. Stability	1	O																					
		9. Perm.	1	O																					
Behaviour Problems	24	BITSEA-P	1	P																					
		CBCL-E	7	P	P			P						P		P		P						P	
		CBCL-SL	1	P																					
		CBCL-Total	2											P	P										
		ECBI-P	1							P															
		Op/Agg	1																						T
		PDR	4		P							P	P	P											
		PDR-E	1																						P
		SAS-P	1																						
		SAS-V	1								P														
		SDQ-H	1																				P		
		SESBI-R	1																					T	
		TRF-E	1																				T		
		TRF-T	1																				T		
Delinquent Behaviour	6	Alcohol attitude	1																					C	
		Antisocial attitude	1																					C	
		Dev. Peer attitude	1																					C	
		Risky Sex. Beh.	1																					C	
		SRD	1																					C	
		Subs. Use	1																						C
Psychological Functioning	18	CBCL-I	5	P		P			P					P										P	
		CBCL-OT	1	P																					
		ECBI-I	3							P													P	P	
		PDR-I	1																					P	
		RAD	1						P																
		RSES	1						C																
		SDQ-C	1						C																
		SDQ-P	2						P													P			
		SDQ-T	1						T																
		SPCC	1																						C
TRF-I	1																						T		
Cog.Function	1	DCCS	1		O																				
Academic	1	Early literacy	1																				O		
Self-regulation	5	BRS-ER	1	O																					
		Cortisol	2		O				O																
		HPA-Axis	1		O																				
		Self-reg composite	1																					P	
Social Competence	10	BITSEA-SC	1	P																					
		BRS-O	1	O																					
		CBCL-SC	1																					P	
		Em.Und.	1																					P	
		PDR-PS	1																					P	
		PIPPS	1																					O	
		SIQ-C	1																						
		SIQ-N	1																					O	
		SIQ-P	1																					O	
		TOM	1		O																				

Table 2 (continued)

Domain	Dom. Freq	Outcome measures	Rep. Freq.	PFR	ABC1	ABC2	FCCT	FFI	PCIT	PSB	KEEP1	KEEP2	KEEP3	PMTO	CBT-PT	CEBPT	IV	IV+CP	KITS	MSS	
Attachment	2	PAD	1		P																
		TAS45	1	O																	
Parent-Child Relationship	7	EAS-S	1				O														
		EAS-ST	1				O														
		EAS-N	1				O														
		EAS-R	1				O														
		EAS-I	1				O														
		IPCI-E	1	O																	
		IPCI-PS	1	O																	
Parental Sensitivity	2	MS	1		O																
		NCATS	1	O																	
Parental Psychological Functioning	11	BDI -II	1															P			
		PED	1															P			
		PSI/NOSI-R	5	P		P		P	P					P							
		PDR-S	1										P								
		PSI-DC	1						P												
		PSI-DI	1						P												
		PSI-PD	1						P												
Parental Skills	19	CAPI	1		P																
		CC	1						P												
		DPICS-II-LP	1						O												
		DPICS-II-neg	1						O												
		DPICS-II-NT	1						O												
		DPICS-II-pos	1						O												
		PBQ-A	1											P							
		PBQ-D	1											P							
		PBQ-E	1											P							
		PBQ-R	1											P							
		PBQ-S	1											P							
		PBQ-W	1											P							
		PPI-AD	1																	P	
		PPI-CE	1																	P	
		PPI-HD	1																	P	
		PPI-PD	1																	P	
		PROP-REINF	1									P									
		PS	2														P	P			
		Parental Knowledge	2	KBPAC	1												P				
RAB	1			P																	
Commitment	1	TIMB	1	P																	
Co-parenting	4	CoP	1																	P	
		CoP-F	1																	P	
		CoP-SS	1																	P	
		CoP-PS	1																	P	
Total outcomes reported per study			122	17	7	4	5	7	12	6	3	4	2	13	3	4	5	11	10	9	
Total self-report measures			87	9	2	4	5	1	8	3	1	2	2	13	2	3	5	11	8	8	
Total clinician assessed (O)			35	8	5	0	0	6	4	3	2	2	0	0	1	1	0	0	2	1	
Total unique measures used			100																		

KEY

<span style="background-color: #008000; color: white; padding: 2px;"> </span> Pre-registered measure	<span style="border: 1px solid black; padding: 2px;">T</span> Teacher report
<span style="border: 1px solid black; border-radius: 50%; padding: 2px;">O</span> Clinician assessed	<span style="border: 1px solid black; padding: 2px;">C</span> Child report
<span style="border: 1px solid black; padding: 2px;">P</span> Parent report	

**Table 2** (continued)

Placement: carer dropout, foster carer dropout rate; Succ Days, number of successful days in foster care; Changes, number of placement changes; Pos Exits, number of positive placement exits; Neg Exits, number of negative placement exits; Breakdown, number of unplanned placement breakdowns; Disruption, Placement Disruption Rate; Stability, Placement Stability/binary: child remain with study caregiver; Perm., Placement permanency/binary: placement stab + legal discharge to study caregiver. Behaviour problems: BITSEA-P, Brief Infant–Toddler Social and Emotional Assessment/problem behaviour scale; CBCL, Child Behavior Checklist/Total; CBCL-E, Externalising Subscale; CBCL-SL, Sleep problem subscale; ECBI-I, Eyberg Child Behavior Inventory/Intensity subscale; ECBI-P, Problem subscale; Op/Ag, Oppositional and Aggressive Behaviour/Latent variable of: CBCL/Oppositional and Aggressive subscales, Teacher Report Form/Oppositional and Aggressive subscales and Conners Teacher Rating Scale–Revised/Oppositional subscale; PDR, Parent Daily Report Checklist; PDR-E, Externalising subscale; SAS-P, Sibling Aggression Scale/Physical subscale; SAS-V, Sibling Aggression Scale/Verbal subscale; SDQ-H, Strengths and Difficulties Questionnaire/Hyperactivity subscale; SESBI-R, Sutter-Eyberg Student Behavior Inventory–Revised; TRF-E, Teacher Report Form/Externalising subscale; TRF-T, Teacher Report Form/Total Problems subscale. Delinquent Behaviours: Alcohol attitude, Positive attitudes towards alcohol use in third grade; Antisocial attitude, Positive attitudes towards antisocial behaviour in third grade; Dev Peer attitude, Involvement with deviant peers at third grade; Risky Sex. Beh., Health-Risking Sexual Behaviours; SRD, Self-report Delinquency Scale; Subs Use, Substance Use, child self-report (Tobacco, Marijuana, Alcohol). Child Psychological Functioning: CBCL-I, Child Behavior Checklist/Internalising subscale; CBCL-OT, Other problems scale; PDR-I, Parent Daily report/Internalising subscale; RAD, Reactive Attachment Disorder Scale; RSES, Modified Rosenberg Self-Esteem Scale; SDQ-C, Strengths and Difficulties Questionnaire/Child report; SDQ-P, Strengths and Difficulties Questionnaire/Parent report; SDQ-T, Strengths and Difficulties Questionnaire/Teacher report; SPCC, self-perception profile for children, global self-worth subscale; TRF-I, Teacher Report Form/Internalising problems. Cognitive functioning: DCCS, The Dimensional Change Card Sort. Academic Functioning: early literacy skills, composite/latent variable of: Concepts About Print Assessment (CAPS), Dynamic Indicators of Basic Early Literacy Skills and Caregiver rating of literacy skills. Self-Regulation: BRS-ER, Bayley Behavior Rating Scales/Emotion regulation subscale; Cortisol, Diurnal Cortisol/Saliva; HPA axis, HPA functioning (salivary cortisol index) pre- and post-exposure to the strange situation; Self-reg composite, Composite of three subcomponents: Emotion Regulation (composite measure); Behaviour Regulation (composite measure), Inhibitory Control (composite measure). Social Competence: BITSEA-SC, Brief Infant–Toddler Social and Emotional Assessment/Social competence subscale; BRS-O, Bayley Behavior Rating Scales/Orientation subscale; CBCL-SC, Child Behavior Checklist/Social Competence; Em. Und., Emotion Understanding; PDR-PS, Parent Daily Report Checklist/Prosocial subscale; PPIPPS, Preschool Penn Interactive Peer Play Scale; SIQ-C, Sibling Interaction Quality/Conflict subscale; SIQ-N, Negative interaction subscale; SIQ-P, Positive interactions subscale; TOM, Theory of Mind Task. Attachment behaviour: PAD, Parent Attachment Diary; TAS45, Toddler Attachment Sort-45. Parent–Child Relationship: EAS-S, Emotional Availability Scales/Sensitivity subscale; EAS-ST, Structuring subscale; EAS-R, Responsivity subscale; EAS-I, Involvement subscale; IPCI-E, Indicator of Parent–Child Interaction/Child engagement subscale; ICPI-PS, Parent support subscale. Parental Sensitivity: MS, Maternal sensitivity observational play; NCATS, Nursing Child Assessment Teaching Scale (NCATS). Parent Psychological Functioning: BDI-II, Beck Depression Inventory; PED, Profile of Emotional Distress; PSI-NOSI, Parenting Stress Index/Dutch adaptation of Parenting Stress Index; PDR-S, Parent Daily Report Checklist/Parental stress; PSI-DC, Parent Stress Index, difficult child subscale; PSI-DI, PSI-PD, Parent–child dysfunctional interaction subscale; PSI-PD, parental distress subscale. Parental Skills: CAPI, Child Abuse Potential Inventory; CC, Parent Conflict Mediation Conflict Checklist; DPICS-II-LP, Dyadic Parent–Child Interaction Coding System/Labelled praise; DPICS-II-neg, Negative composite; DPICS-II-NT, Negative talk; DPICS-II-pos, positive composite; PBQ-A, Parenting Behavior Questionnaire/Autonomy granting subscale; PBQ-D, Discipline subscale; PBQ-E, Explaining subscale; PBQ-R, Responsiveness subscale; PBQ-S, Strictness subscale; PBQ-W, Warmth subscale; PPI-AD, Parenting Practices Interview/Appropriate discipline subscale; PPI-CE, Clear expectations subscale; PPI-HD, Harsh discipline subscale; PPI-PD, Positive discipline subscale; PROP-REINF, Ratio of positive reinforcement to discipline; PS, Parenting Scale. Parental Knowledge: KBPAC, Knowledge of Behavioural Principles as Applied to Children; RAB, Raising a Baby. Parental Commitment: TIMB, This Is My Baby. Co-parenting relationship: CoP, Co-parenting composite measure; Cop-F, Flexibility subscale; Cop-SS, Mutual social support subscale; Cop—PS, Problem solving subscale

	ABC 1	ABC 2	CBTPT	CEBP PT	FCCT	FFI 1	IV-B	IV+CP 1	KEEP 1	KEEP 2	KEEP 3	KITS 1	MSS 1	PCIT 1	PFR-1	PMTO	PSR 1
Random sequence generation.	?	--	--	--	--	--	?	+	?	?	?	--	--	--	--	--	+
Allocation concealment.	--	?	--	?	--	--	?	?	?	--	?	--	?	--	--	--	?
Performance Bias	--	+	+	+	?	+	+	+	+	+	+	+	+	+	--	+	+
Outcome Assessment Bias	--	+	+	+	--	--	+	?	+	+	?	--	--	--	--	--	--
Attrition Bias	?	--	+	--	?	+	+	--	+	--	--	?	--	--	--	--	--
Reporting Bias	+	?	?	?	--	?	+	?	?	?	+	+	?	--	+	?	?
	--		LOW risk of bias			?	UNCLEAR risk of bias			+	HIGH risk of bias						

**Fig. 5** Risk of bias assessment of included studies

a mean of 7 outcome variables, so that 90 outcome measures were unique across the seventeen trials. Eighty-seven outcomes were self-report measures, and thirty-five were clinician-assessed. Four of the studies accounted for 66% of clinician-rated measures leaving the majority of self-report (82%) measures in the remaining thirteen studies.

Child behaviour was the most commonly reported outcome domain across trials, with fifteen interventions (88%) reporting some kind of behavioural outcome. Despite this convergence, seven different behavioural outcome instruments were used to measure behaviour problems. In addition, a range of subscales of these measures was used (e.g. externalising, aggressive, and oppositional subscales) so that in total 14 different measures or subscales were used across the fifteen trials.

Parental stress and psychological well-being were the next most commonly reported outcome domains across trials, with eight studies (47%) reporting an outcome in both of these domains. The Parenting Stress Index (PSI) (Abidin 1990) was the most commonly used instrument being included in five studies (29%). Three other parenting stress outcome measures were in used to evaluate outcomes this domain (see Table 2).

Outcomes associated with the development of effective parenting skills were also common, reflecting that the aim of many interventions was to improve foster/kinship carers' capacity to manage challenging behaviour of foster children. Five studies reported outcomes in this domain, using seven psychometric instruments composing nineteen different subscales.

Authors provided citations to articles establishing the validity and psychometric properties of outcome variables in the majority of cases (85%). Most of the commonly used scales, however, cited articles that were not normed for the alternative care child population, presenting a potential challenge to their external validity. Some scales were custom-made by developers and evaluators of the intervention (e.g. ABC, KEEP, PFR) or were custom-made composites of several scales used to provide a global measure of a domain (e.g. MSS). Some of these custom-made measures failed to provide information or reference to construct validity or psychometric properties (e.g. Reactive Attachment Disorder in PFR, Parent Attachment Diary in ABC1) or provided links to articles that were unpublished, or inaccessible on journal databases (e.g. Parent Daily Report in KEEP).

### Risk of Bias in Included Studies

Figure 5 provides a visual overview of the risk bias assessment consensus (for detailed of risk of bias judgements comments, see 'Appendix').

### Selection and Allocation Bias

Selection bias can threaten the internal validity of trials by introducing non-systematic differences between intervention and control groups (Higgins et al. 2011). Methods used to minimise selection bias include allocating participants, using random sequence generation and concealing participant allocation from researchers.

All seventeen included studies were described as randomised controlled trials and reported a range of methods used to minimise selection bias. Ten of the seventeen trials reported methods of random sequence generation that met criteria for 'LOW' risk of selection bias. Four did not provide sufficient information about randomisation and were judged 'UNCLEAR'. One was considered 'UNCLEAR' because, despite a well-reported randomisation procedure, there were significant unexplained differences in baseline levels of the primary outcome measure (KEEP3). Two studies received a 'HIGH' risk of bias rating due to a failure to provide sufficient information indicating randomisation had taken place together with significant baseline differences that may favour the intervention group (KEEP1, PSB).

Nine of the seventeen trials were judged to have 'LOW' risk of bias in the allocation bias domain because they described methods of allocation concealment sufficient to minimise risk of selection bias. Eight trials did not provide sufficient information about allocation concealment to make a clear judgment, and received an 'UNCLEAR' rating.

Reporting demographic data independently for the intervention can help to ensure randomisation was successful and trial groups are equivalent (Higgins et al. 2011), as unreported differences at baseline can lead to biased estimates of between-group differences (Clarke and Oxman 2000). Eight studies (47%) failed to report any baseline demographic data independently for intervention and control (ABC2, CEBPT, CBT-P, IY + CP, KEEP1, PCIT, KITS, MSS). The remaining studies reported a mixed range of characteristics for control and intervention groups (see Fig. 3). Because placement history, maltreatment history, foster carer experience and family context are all well-established predictors of foster child difficulties and outcomes, ensuring group equivalence on these factors is an important guard against potential bias. These standards of reporting were reflected in the Cochrane risk of bias ratings given to the studies.

### Performance Bias

Performance bias refers to participants' exposure to factors other than those related to the intervention. To achieve a 'LOW' risk of performance bias, participants and researchers a required to be blind to treatment condition, so-called double blinding. Blinding is a recognised challenge in psychosocial interventions, as participant blinding requires

an active control group (Goldbeck and Vitiello 2011). For this reason, only two studies reported double blinding and a ‘LOW’ rating for performance bias. One study informed participants about blinding, to ensure that 11 foster children were not aware of treatment condition; this study received an ‘UNKNOWN’ rating for performance bias. Fourteen studies received a ‘HIGH’ judgment for risk of performance bias because they used a treatment-as-usual control group.

### Detection Bias

Detection bias can result when the assessment of outcomes changes in response to knowledge of participants’ allocation status. One author (AD) was of the opinion that, as parent self-reports were used for outcome assessment, parents were in all senses the assessors of the study outcome. Overall, nine studies were judged to have a ‘LOW’ risk of detection bias, because at least some of the assessments were conducted by blinded research assessors. Two studies were judged to have ‘UNCLEAR’ risk of detection bias, because researchers were reported as blind but only self-report measures were used. Six studies received a judgement of ‘HIGH’ risk of detection bias, because they used interview assessments in which neither party were reported as blinded.

### Attrition Bias

Attrition bias refers to the systematic differences between groups that may bias analysis of post-intervention outcomes (Higgins et al. 2011). Ten of the trials were judged to have ‘LOW’ risk of attrition bias, because attrition rates were less than 20%, and measures were taken to analyse outcome data to compensate for drop-outs using Last Outcome Carried Forward (LOCF) analysis. Three studies received an

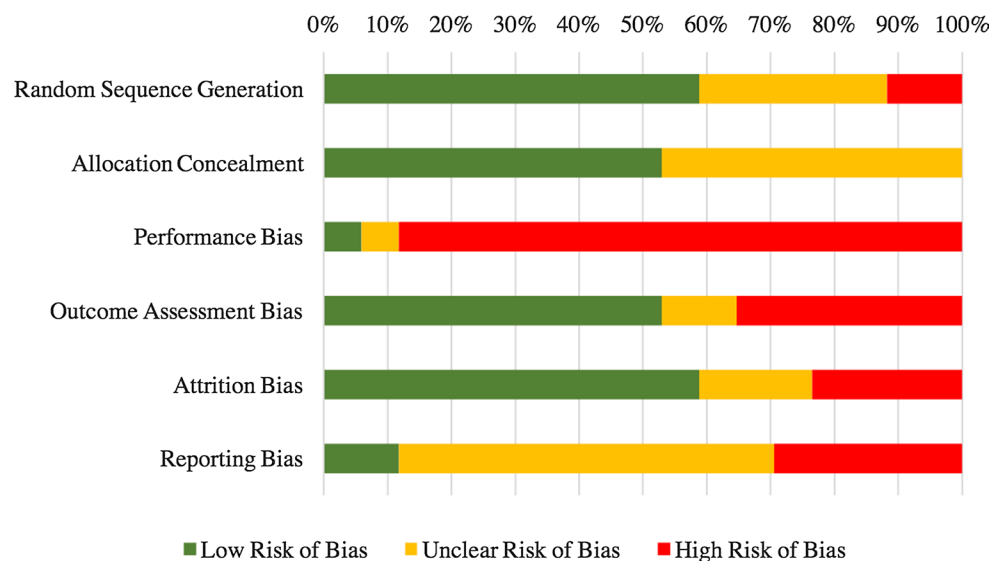
‘UNCLEAR’ risk of bias, because they did not provide adequate information to provide sufficient information about attrition rates in the study. Four studies were judged to be at a of ‘HIGH’ risk of bias because they had levels of attrition, included no information about attrition at all and did not report the use an intention to treat analysis to compensate for drop-outs.

### Reporting Bias

Reporting bias refers to systematic differences between reported and unreported findings and includes the selective reporting of both trials and trial outcomes (Higgins et al. 2011). Two of the included interventions (18%) were judged to have a ‘LOW’ risk of bias because they pre-registered both the aim of the intervention and the outcome measures that were subsequently reported (FCCT, PFR). Ten of the included interventions received an ‘UNCLEAR’ judgement of reporting bias because the trial was not pre-registered and authors provided no evidence that all of the outcome measures used in the trial were reported. Finally, five of the included trials were judged to be at HIGH risk of bias. Of these, one trial received this rating because authors pre-registered outcome measures that were not subsequently reported and also reported outcomes that were not registered (MSS). Remaining studies received this rating because the primary outcome measure reported in the abstract was the only significant finding of many reported outcomes (IY), reported a primary outcome in a domain not associated with the intervention aims (PMTO) or added to the trial registry shortly before a report was submitted for publication, but after the trial had been completed (ABC1).

Trial registration is now seen as an essential safeguard against publication bias in scientific research (DeAngelis

**Fig. 6** Summary of risk of bias across included trials



et al. 2004). Only five of the seventeen trials (29%) registered outcome variables in a trial registry (ABC1, KEEP3, PFR, KITS, MSS) (for details see ‘Appendix’). Of these, one trial added outcome measures into the registry after the trial completion (ABC1) and two registered only outcome domains, rather than specific assessment instruments (MSS, KITS). Only two interventions (11%) pre-registered the specific outcome instruments that were subsequently used as evidence for programme efficacy in published articles (PFR, KEEP3). These limitations were reflected in the ratings of reported bias given to included trials.

### Interacting Bias Across Domains

Assessing risk of bias is not algorithmic or automatic process, but depends on assessor judgement based on a detailed analysis of reporting (Higgins et al. 2011). In some studies, the interactions between bias across domains led to unclear or misleading reporting of results. Specifically, reporting bias, combined with randomisation, allocation or attrition bias led to the reporting of intervention efficacy that was not reflected in systematic differences between intervention and control groups. One randomised study reported significant benefits to foster parents and children in the intervention (KEEP1). However, the significantly greater reduction in behaviour problems in the intervention group (e.g. different in change scores) compared to control was due to differences at baseline, with no significant between-group difference in post-intervention scores. Another randomised study reported significantly larger changes in mean behaviour problems for the intervention group, compared to control (IY). However, this larger change was not reflected in lower levels of post-intervention problem behaviours for the intervention group. Instead, reported changes resulted from higher baseline levels of behaviour problems in the intervention group, suggesting problems with randomisation. Additionally, authors reported positive and significant effect sizes for the intervention based solely on pre–post-scores (paired *t* tests), rather than the negative effect sizes associated with between-group differences in post-intervention scores. A third study clearly reported potential randomisation problems that had failed to result in equivalence between control and intervention groups (KEEP3). However, authors also reported an HLM analysis that suggested significantly larger decreases in behaviour problems were found in the intervention group, compared to control. These differences in change scores, reported as a significant group  $\times$  interaction effect, were the result of baseline differences in behaviour problems. Even though post-intervention scores were equivalent, suggesting there was no benefit to participants randomised to the intervention, significant and positive within-group effect sizes were reported. Finally, the significant findings reported were based on focal siblings swapped in during the intervention

due to attrition. An additional analysis that included the same participants pre- and post-intervention was not significant, even for within-group changes. This suggests that potential allocation or attrition bias may have contributed to reporting that did not reflect intervention effects on focal children.

### Summary of Risk of Bias Assessment

Figure 6 summarises the mixed study quality. Fewer than fifty per cent of trials reported both the randomisation and allocation concealment procedures that are required for a low rating of selection bias. Because of the almost universal dependence on self-report measures in single blinded trials, only two of the seventeen included studies had low risk of bias of both performance and detection bias (ABC, PFR), although this reflects a shortcoming common in evaluations of psychosocial interventions that depend on self-report. Across the included studies, risk associated with attrition bias was managed best, with almost 60% reporting low attrition and the use of analytic methods to compensate for participant drop-outs. In contrast, over 80% of included studies had a low or unclear level of reporting bias. This reflected the low level of trial registration, with only five of the seventeen trials using pre-registration. Of those registered trials, only one trial had a complete concordance between registered and reported outcomes (KEEP2; see ‘Appendix’).

## Discussion

This review aimed to provide a systematic analysis of the methodological factors that threaten the internal and external validity of interventions for foster and kinship carers and an overview of the clinical heterogeneity currently limiting the meaningful synthesis and comparison of trial results.

The results of this analysis show how the complex and diverse context of the field (see, for example, Jones et al. 2011) gives rise to a series of methodological challenges to robustly evaluating interventions for children in alternative care. The results mirror comments by Tarren-Sweeney (2013a, b, 2014) in which he outlines the range of challenges facing researchers seeking to establish a robust evidence base for children with complex trauma, attachment problems and children in alternative care. Tarren-Sweeney identifies continuing tensions between the principles of working from an empirical evidence base and the enormous complexity of human psychological development and morbidity, much of which is yet to be adequately measured or conceptualised.

Researchers in the field of child protection have spent over 40 years attempting to make sense of this complexity, using the understanding garnered from research to develop a range of interventions that aim to improve foster/kinship care



in a meaningful way. Responding to the increasing demands of evidence-based practice, researchers have subjected their interventions to increasingly sophisticated trials that seek to establish their efficacy and effectiveness. Yet despite improvements in methodology over time, there continue to be several challenges in adapting the diverse and complex interventions used in family-based alternative care to the rigid requirements of randomised trials. The full extent of this complexity and its impact on the clinical heterogeneity, internal validity and external validity of the included studies are reflected in the results of this review.

### External Validity

There were some limits to external validity that may restrict the generalisability of trial results. In contrast to previous reviews that noted several non-validated outcome variables were used to evaluate trials, results of this review found that the great majority were validated. This difference may relate to inclusion criteria that limited this review to high-quality randomised trials. Something that previous reviewers rarely commented on was the use of validated measures that were normed for non-foster/kinship care populations, a factor that may present a threat to the external validity of the trials (Locke and Prinz 2002). Unfortunately, the measures that were developed specifically for use in the trials were those that did not provide information about validation or norming with children in care.

The clinical diversity of children in care makes the reporting of demographic information particularly important for establishing the external validity, especially given that children in care at varying developmental stages, with different maltreatment histories, placement histories and psychological and relational problems have different needs (Schuengel et al. 2009). Without accurate reporting of such information, there is no way of knowing which carers and children will benefit from an intervention with demonstrated efficacy. Fortunately, the reporting of participant characteristics was generally good, with some studies collecting a range of detailed information. Even in these cases, however, this may still not be sufficient to guarantee that trials can be generalised to clinical populations. Complex and costly interventions need to be provided to those who will benefit from them most, and those that need them most must be shown to receive benefit from them (Chambless and Ollendick 2001). If trials allocate participants that are not screened for the problems the intervention is designed to target, then the results of that trial cannot be said to demonstrate effectiveness that population. Thus, the trial's external validity may be compromised (Rothwell 2005). Of the included interventions, only four (ABC2, PCIT, PMTO, CEBPT) used participants assessed for difficulties that were reflected in their programme's

treatment aims and could, as a result, demonstrate strong external clinical validity.

### Clinical Heterogeneity

The purpose of systematic reviews and meta-analyses of clinical trials is to synthesise and compare the results of trials to provide an overall summary of intervention effectiveness (Haidich 2010). Reviewers of foster/kinship carer interventions face particular challenges in completing this task. In contrast to trial populations with specific, homogenous problems, children in care face multiple challenges and have multiple needs manifested across behavioural, developmental, psychological, physiological and cognitive domains. In turn, these challenges have further implications for the child, the foster/kinship carer, their relationship and the stability of the placement (Jones et al. 2011; Oosterman et al. 2007).

The results of this review suggest that the extent of clinical heterogeneity in participants, aims settings and outcomes limit the meaningful synthesis of these interventions. It is unlikely, for example, that the effectiveness of an attachment intervention on improving HPA axis functioning in infants can be meaningfully compared to a behavioural intervention targeting teenage delinquency. The enormous range of outcomes used to evaluate these different interventions, and the lack of coherence and consensus over which instruments or scales should be used to determine efficacy, provides a further barrier to comparing or synthesising these results, limiting both their value in informing practice and policy and their capacity to inform future research.

More positively, the results of this study indicate that research is now beginning to focus on developing different interventions for specific developmental stages, with attachment-based interventions (ABC1, ABC2, FFI, PFR) focusing on encouraging warm, responsive care that increases attachment security for children under the age of six. In contrast, Social Learning Theory approaches (IY, PMTO, CEBPT, CBT-PT, KEEP2 and KEEP3) tend to include older participants and aim to modify the coercive cycle of negative reinforcement that can result from child behaviour problems. Similarly, MSS and KITS interventions were targeted at children at a particular developmental stage, and also show that efforts are being made to link interventions to identified needs. Other approaches not included in this review (e.g. Multi-Dimensional Foster Care, MTFC; Fisher et al. 2005; Fisher and Kim 2007) are showing promise in targeting interventions towards treatment need.

As such, the results of this review suggest that that clinical heterogeneity—the diversity of interventions, participants and outcomes—may present real challenges to those who wish to synthesise and compare interventions in a meaningful way. Existing reviews suggest a range of research questions and hypotheses that can be tested in

future interventions. These interventions can be targeted at participants screened for specific problems, at specific developmental stages using a unified set of outcome variables would help to address these challenges and provide material more amenable to robust meta-analyses and systematic reviews.

### Challenges to Internal Validity

Randomised controlled trials—and the statistical methods used to analyse them—are designed to test a hypothesis: that a particular intervention is effective in treating a particular disorder or problem, in a specified population (Kendall 2003). Ideally, an appropriate, validated and pre-specified primary outcome measure (or composite measure) is used to detect differences in treatment and control groups. The null hypothesis (that the intervention has no effect) is rejected if the outcome is sufficiently unlikely to be a result of chance (e.g.  $p < 0.05$ ). Departures from this approach challenge both the internal validity necessary to ensure efficacy and the external validity that supports generalisation of trial results. Differences between the trials' hypotheses, interventions, setting and participants challenge the ability of reviewers to meaningfully synthesise and compare the efficacy of interventions.

In the included studies, the diverse needs of participants, the complexity of the interventions and the measurement of effect using multiple outcomes over a number of domains potentially challenges unbiased trial evaluation. Firstly, hypothesis testing depends on a clear link between theory, intervention aim, hypothesis and outcome measures. Where a hypothesis is too vague: 'this intervention may be effective in one of a broad range of outcomes' the statistical principles of null hypothesis testing first developed by Fisher and Neyman are violated (Simonsohn et al. 2014) and rejecting the null hypothesis no longer signifies a true treatment effect (Frick 1996). Ideally then, clinical trials should be confirmatory and depend on a priori hypotheses that are generated by theory and exploratory research and then tested by the experimental research. In contrast to these requirements only two of the seventeen included (KEEP3, PFR) studies used pre-registered a priori outcomes and only five trials reported a primary outcome post hoc, with fourteen trials used multiple measures that were not specified before the trial.

There is an additional issue associated with the use of multiple outcomes without a specified primary outcome: the selective reporting of significant outcomes that can lead to a biased evaluation of efficacy. This is a potential problem that even more recent methods such as  $p$ -curve analysis (funnel plot, variance testing) nor meta-analytic methods are able to detect (Bishop and Thompson 2016; Ioannidis et al. 2008). It was also notable that only one study reported adjustments to  $p$  values to compensate for the reporting of

multiple outcomes, despite the established  $p$ -inflation that results in these cases (PMTO). As a result, some trial results near the critical value of  $p = 0.05$  that were reported as significant may be not be robust treatment effects (Feise 2002).

Finally, although it makes sense for interventions that address complex problems to have several aims (something that is common in many psychosocial interventions), the aims themselves also need to be pre-registered. Unfortunately, results indicate that some included studies published multiple reports of the same trial, but modified and adapted the stated aims of the intervention consistent with the outcomes that were reported. This practice again risks violating the principles of hypothesis testing, challenges the theoretical coherence of the intervention and makes statistical tests less reliable in detecting a true effect (Frick 1996). As for external validity, targeting interventions for participants screened for specific needs and using single primary outcomes to measure effectiveness will likely increase effect, which is additionally important given the identified lack of power in many of the included studies.

The intense pressure to demonstrate effects in research (Nissen et al. 2016) using RCTs (Ioannidis 2005) has led to an increased attention on research guidelines that ensure high standards of trial methodology and reporting (Higgins et al. 2011) and an increase in the overall quality of RCTs (Begley and Ioannidis 2015). In general, however, the studies included in this review tended to rely on self-report data and this was important in all but one of the included trials receiving a rating of high risk of bias. In addition, the use of non-active control groups means that participants cannot be blinded to group status. Given that these participants also played a central role in assessing child outcomes (e.g. using CBCL, PDR), the combination of performance and detection bias presented a significant risk to study results. Even leaving aside performance and reporting bias domains aside, only one of the included trials (PFR) had an overall low risk of bias rating using the Cochrane Handbook of Standards for RCTs. It is notable that this trial detected only one significant change in the 16 outcomes related to interventions aims. As such, trial results of the included interventions that indicate treatment efficacy must be interpreted with some caution.

### Limitations

The clinical heterogeneity in the interventions reviewed here extends to the entire field of alternate care that lies beyond the scope of this review. 'Alternative care' is a general term for children with differing legal status, under varying forms of care (e.g. kinship vs. foster care), in a variety of settings (e.g. residential vs. family care). For practical reasons, this review focussed on interventions targeted towards foster and kinship carers, excluding several important interventions for

children who have experienced maltreatment and separation from their birth parents. As a result, several targeted and robust trials of interventions (e.g. Multi-dimensional Foster Care for Preschoolers; MTFC-P) were excluded, as were RCT evaluations that were published only in white papers, such as the Fostering Changes Programme, which has robust methods and promising results (Briskman et al. 2012). For this reason, this methodological review cannot be taken as indicative of all research in alternative care, or as providing a complete picture of the field. A full review of the field, conducted by a large, well-funded team of researchers (for example, see Fraser et al. 2013) would provide a more comprehensive critique and therefore a more informative review of methods used across the entire field of alternative care.

A further limitation was that characteristics of trial participants were not compared with norms of the general foster care populations. Assessing whether the trial participants were representative of the wider population of carers and children in care, in the outcome measures used to determine effectiveness, would have provided a more systematic analysis of the external validity. Such an analysis was precluded by the heterogeneity of the population, of the outcomes used to establish norms, and the lack of population-wide data.

## Conclusions

The complex challenges of the field continue to present challenges to current standards of trial methodology or reporting, with the result that evidence of the effectiveness of interventions for foster/kinship carers and children remains limited. The field is yet to develop explicit hypotheses concerning which interventions are effective in treating which difficulties for which participants, and then test those hypotheses with interventions targeted at those problems using robust methods and appropriate outcome measures.

## Recommendations

In line with previous reviewers (e.g. Dorsey et al. 2008; Everson-Hock et al. 2012; Festinger and Baker 2013; Kerr and Cossar 2014; Kinsey and Schlösser 2013; Luke et al. 2014; Tarren-Sweeney 2014), the results of this review suggest further research is needed to provide robust evidence of effectiveness in the context of the complex needs and difficulties of foster/kinship families. This context is the focus of an enormous and growing body of research and provides an opportunity to align evaluation trials to current research,

and test focussed hypotheses and research questions in controlled conditions.

In order to test these hypotheses and provide robust evidence that interventions are effective, a clear link needs to be drawn between the identified needs of a specific population (e.g. age, presenting problem), the aims of the intervention and the outcome measures used to detect whether those needs are met. If possible, the use or development of appropriate measures, normed and validated for the population would help to improve their external validity (Tarren-Sweeney 2014). Evidence-based practice is about the judicious application of evidence to the individual patient (Sackett et al. 1996). If evaluation trials are to inform clinicians, they need to know for whom they are effective (Tarren-Sweeney 2014).

Participant characteristics and needs, intervention aims and outcome measures should be registered a priori, and trials should follow all recommended Cochrane procedures to minimise bias. The robust results from such trials would do much to clarify the significant uncertainty that remains in the field. Outcomes need not be limited to single primary outcomes. Additional aims, outcomes and mediation effects can be included into research trials to generate further exploratory hypotheses for confirmatory tests in subsequent RCTs, further extending the research field. Clinical heterogeneity between studies needs to be assessed prior to any meta-analytic or systematic synthesis and precludes an assessment of programme effectiveness (Ioannidis 2005). Reviewers seeking to make sense of this complex field would benefit from narrowing the scope of their reviews to particular population sub-groups or participant needs, approaches similar to existing reviews of Attachment-Based (Kerr and Cossar 2014) and Cognitive Behavioural (Turner et al. 2007) interventions. Given the consistent improvements in trial methodology over time, there is promise that future research will help to unravel the daunting complexity of the field and enable more robust evaluations of interventions for foster families.

**Funding** Funding was provided by the University of Technology Sydney.

## Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical Approval** Ethics approval was obtained from the University of Technology Sydney Human Research Ethics Committee.

## Appendix: Characteristics of Included Studies and Risk of Bias Ratings

### Characteristics of ABC1: attachment and biobehavioral catch-up intervention

Published article/s	Bick and Dozier (2013), Dozier et al. (2009), Dozier et al. (2008), Dozier et al. (2006) and Lewis-Morrarty et al. (2012)	
Methods	Preregistered RCT; NIH clinical trial registry. Clinicaltrials.gov Identifier: NCT00056303	
Setting	USA	
Inclusion criteria	Foster carers with children aged 15–24 months, referred for relational interventions	
Participant characteristics	93 Foster Children aged 15–24 months ( $M = 19.7$ ) and 97 foster carers referred for relational interventions	
Intervention/control*	ABC intervention	Developmental education for families (DEF)
Group ( $n$ )	$n = 46$ (Dozier et al. 2008)	$n = 47$ (Dozier et al. 2008)
Duration	10 × 1 h weekly sessions	10 × 1 h weekly sessions
Baseline demographics*	Mean age 20.0 months 59% female Ethnicity: 81% AA, 2% Hm 17% W	Mean age 19.5 months 43% female Ethnicity: 66% AA, 5% H, 29% W
Reported outcomes	Placement history ( $N$ ): 64% placed < 1 month, 24% placed 1–12 months, 10.8% placed 15–36 months	
	Maltreatment history ( $N$ ): All placed due to neglect OR parent psychopathology OR incarceration	
	Behavioural problems: NR Foster carer experience: NR Kinship versus non-kinship: NR	Behavioural problems: NR Foster carer experience: NR Kinship versus non-kinship: NR
	Attachment behaviour: Parent Attachment Diary (PAD) (M. Dozier and Stovall 1997) Self-regulation: Diurnal Cortisol; HPA functioning pre and post exposure to the strange situation Behaviour problems: Parent Daily report (PDR) (Chamberlain and Reid 1987) Social competence: Theory of Mind Task (TOM) (Oswald and Ollendick 1989) Cognitive functioning: The Dimensional Change Card Sort (DCCS) (Zelazo 2006) Parental sensitivity: Observational play task (MS) (Bick and Dozier 2013)	
Registered primary outcome(s)	No primary outcomes registered before intervention	
Reported aim(s) of intervention	Treatment for the effects of relationship disruption and improve behavioural, emotional and neuroendocrine dysregulation	
Reported aim(s) of studies	To assess the effectiveness of ABC regarding HPA functioning, attachment behaviours, sensitive caregiving, cognitive flexibility and theory of mind skills	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation Bias	Unclear	Randomisation method was not clearly reported
Allocation Bias	Low	Reported as double blinded study in trial registry
Performance Bias	Low	Participants blind to condition, active control used
Detection Bias	Low	Clinician assessment reported as blinded
Attrition Bias	Unclear	Attrition rates vary from cascading allocation of participants, with analysis at several time points
Outcome Bias	High	Trial pre-registered, but outcomes unregistered until post trial. Multiple outcome variables used, with successive study waves permitting outcome selection

### Characteristics of ABC2: attachment and biobehavioral catch-up intervention

Published article/s	(Sprang 2009)
Methods	Unregistered RCT
Setting	USA
Inclusion criteria	Children < 6 months, diagnosed with attachment-related problems that threaten to disrupt placement, without severe mental illness, 0–6 years

Participant characteristics	<i>N</i> = 58 foster and adoptive children aged 0–6 years ( <i>M</i> = 3.5 years; <i>SD</i> = 2.6 years)	
	<i>N</i> = 58 foster carers and adoptive parents	
Intervention	ABC intervention	TAU
Group <i>n</i>	<i>n</i> = 29	<i>n</i> = 29
Baseline demographics ( <i>N</i> )	Mean age: 42.5 months Gender: 49% female Ethnicity: 85% ‘white’ Placement history: NR Maltreatment history: NR Behavioural problems: NR Foster carer experience: NR Kinship versus non-kinship: NR	
Duration of study	10 × 1 h weekly sessions	
Reported outcomes	Behaviour Problems: Child Behaviour Checklist-Externalising subscale (CBCL-E) (Achenbach and Rescorla 2001) Psychological Functioning: Child Behaviour Checklist-Internalising subscale (CBCL-I) (Achenbach and Rescorla 2001; Koren-Karie and Markman-Gefen 2016) Parental Skills: Child Abuse Potential Inventory (CAPI) (Milner 1990) Parent Psychological Functioning: Parenting Stress Index (PSI) (Abidin 1995)	
Reported aim(s) of intervention	This manualised intervention helps caregivers learn optimal sensitive parenting behaviour	
Reported aim(s) of studies	The objective of this study was to assess the effects of participation in a relational intervention program on child abuse potential, parenting stress and child behaviour in maltreated children	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	Randomised with a fixed randomisation procedure with allocation of 1 of every four to intervention group
Allocation bias	Unclear	Allocation sequence not described in detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU control group used. No participant blinding possible
Detection bias	High	No reported blinding of clinician assessors; non-blinded participant assessment (self-report)
Attrition bias	Low	< 20% attrition rate; ITT analysis conducted.
Reporting bias	Unclear	No pre-registration of trial protocol or outcome variables was conducted.

Characteristics of CBT-PT: cognitive behavioural parent training program

Published article/s	Macdonald and Turner (2005)	
Methods	Unregistered RCT	
Setting	United Kingdom	
Inclusion criteria	Foster carers in local area providing full-time care	
Participant characteristics	(mean age of child not reported); <i>N</i> = 117 foster carers and adoptive parents with children in long-term foster care	
Intervention	CBT-PT Intervention	Wait List Control

Group <i>n</i>	<i>n</i> = 67	<i>n</i> = 50
Baseline demographics ( <i>N</i> )	Mean age: NR % female: NR Ethnicity: NR Placement history: NR Maltreatment history: NR Behavioural problems: NR Foster carer experience: <i>M</i> = 8 years Kinship versus non-kinship: NR	
Duration of study	4 × 5 h weekly sessions	
Reported outcomes	Behaviour Problems: Child Behavior Checklist—Total (CBCL) (Achenbach and Rescorla 2001) Parental Knowledge: Knowledge of Behavioural Principles as Applied to Children (KBPA) (O'Dell et al. 1979) Placement: Number of unplanned placement breakdowns	
Reported aim(s) of intervention	Increase ... knowledge of behavioural problems and how to manage them (in cognitive-behavioural terms), (ii) increase in skills in the management of behaviour problems, (iii) decrease in behavioural problems (or their frequency or severity), and (iv) increase in confidence in carers' perceived ability to cope with/manage children with behavioural problems	
Reported aim(s) of studies	As above	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	Random numbers table used for allocation
Allocation bias	Low	Random numbers table used for allocation
Performance bias	High	No blinding reported. TAU control group used. No participant blinding possible
Detection bias	High	No reported blinding of clinician assessors; non-blinded participant assessment (self-report)
Attrition bias	High	> 20% attrition rate; no ITT analysis conducted, no explanation for drop outs. No CONSORT diagram provided
Reporting bias	Unclear	No pre-registration of trial protocol or outcome variables was conducted

#### Characteristics of CEBPT: cognitive behavioural parent training program

Published article/s	(Gavita et al. 2012)	
Methods	Unregistered RCT	
Setting	Romania	
Inclusion criteria	Foster carers and adoptive parents of children aged 5–18 with Child Externalising disorders above 93 <sup>rd</sup> percentile (CBCL) kinship carers excluded	
Participant characteristics	<i>N</i> = 97 foster children 5–18 years ( <i>M</i> = 9.5 years; <i>SD</i> = 3.47 years); <i>N</i> = 97 non-kin foster carers and adoptive parents	
Intervention	CEBPT intervention	Wait list control
Group <i>n</i>	<i>n</i> = 56	<i>n</i> = 41
Baseline demographics of foster children ( <i>N</i> )	Mean age: 9.5 years ( <i>SD</i> 3.47) % female: NR Ethnicity: NR Placement history: NR Maltreatment history: NR Behavioural problems: NR Foster carer experience: NR Kinship versus non-kinship: 0/100%	
Duration of study	4 × 4 h weekly sessions + 4 h follow-up	



Reported outcomes	Behaviour Problems: Child Behavior Checklist-Externalising Subscale (CBCL-E) (Achenbach and Rescorla 2001) Parental Skills: Parenting Scale (PS) (Arnold et al. 1993) Parent Psychological Functioning: Profile of Emotional Distress (PED) (Opris and Macavei 2007) Placement: Placement Disruption Rate (%)
Reported aim (s) of intervention	Reduce child externalizing behaviour disorders by focusing first on teaching foster parents functional emotion-regulation strategies and by supporting better implementation of parenting skills
Reported aim (s) of study(s)	To investigate the efficacy of a short enhanced cognitive-behavioural group parenting program for reducing externalizing behavior disorders in Romanian foster children, increasing placement stability, and reducing foster parents' emotional distress

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	Reported a block randomisation process
Allocation bias	Unclear	Allocation sequence not described in detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU control group used. No participant blinding possible
Detection bias	High	CBCL + PS: No reported blinding of clinician assessors; non-blinded participant assessment (self-report) Placement Breakdown: No researcher blinding reported
Attrition bias	Low	Overall attrition < 20%; ITT analysis reported
Reporting bias	Unclear	No pre-registration of trial protocol or outcome variables was conducted

Characteristics of FCCT: foster carers communication training

Published article/s	Minnis and Devine (2001)	
Methods	Unregistered RCT	
Setting	United Kingdom	
Inclusion criteria	Foster carers and children in Scottish Council areas, aged 5–16	
Participant characteristics	N = 121 foster families; 181 foster children aged 5–16 (M = 11.25 years)	
Intervention	FCCT intervention	TAU
Group n	n = 80	n = 80
Baseline demographics	Mean age: 10.9 years	Mean age: 11.6 years
	% female: 42	% female: 44
	Ethnicity: NR	Ethnicity: NR
	Placement history:	Placement history:
	Mean time w/current carer: 29 months	Mean time w/current carer: 32 months
	Maltreatment history: % abused 80, neglected, 76, abused and neglected 89	Maltreatment history: % abused 87; neglected 72; abused and neglected 93
	Behavioural problems: % classified as psychiatric cases, % in residential care	Behavioural problems: % classified as psychiatric cases, % in residential care
Foster carer experience: Median (range); no. children prev. cared for: 18 (1, 91)	Foster carer experience: Median (range) no. children prev. cared for: 14 (1, 140)	
Kinship versus non-kinship: NR	Kinship versus non-kinship: NR	
Duration of study	2 × 6 h sessions, consecutive days, 1 × 6 h, 1 weeks later	
Reported outcomes	Psychological Functioning: Strengths and Difficulties Questionnaire—child, parent, teacher report (SDQ-C/P/T) (Goodman 2001); Reactive Attachment Disorder Scale (RAD) (Millward et al. 2006); Modified Rosenberg Self-Esteem Scale (RSES)	
Reported aim (s) of intervention	To help foster carers develop their skills in communication and increase their confidence in their ability to cope with their foster child's feelings and behaviour	
Reported aim (s) of study(s)	(to answer the following questions:) What emotional and behavioural problems do children have in foster care? Can training help foster carers to reduce children's problems?	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation Bias	Low	Randomised by independent personnel. no significant differences in demographic data of groups. Analyses adjusted to compensate for baseline differences
Allocation Bias	Low	Families were randomised, after consent, using random permuted blocks of block size. The trial was single blind. The randomisation list and identifying information such as receipts were handled only by the study secretary
Performance Bias	Unclear	Single blinding reported, carers informed about blinding; outcome measures from parent, teacher and carer used
Detection Bias	Low	Multiple non-participant assessors blinded to group
Attrition Bias	Unclear	54% attrition for main immediate outcome variable 20% for 9 month follow-up
Reporting Bias	Low	Trial not registered. None of the results were significant

#### Characteristics of FFI: foster-carer foster-child intervention

Published article/s	Van Anandel et al. (2016)	
Methods	Unregistered RCT	
Setting	The Netherlands, In-home visits	
Inclusion criteria	Foster carers with expected placement duration > 6 months, excluding children with birth defects, severe cognitive dysfunction, and problems indicating treatment	
Participant characteristics	N = 123 Foster and Kinship Carers; N = 123 foster children aged 0–5	
Intervention characteristics	FFI intervention	CAU
Group <i>n</i>	<i>n</i> = 65	<i>n</i> = 58
Baseline demographics	Mean age: 1.6 years % female: 51 Ethnicity: NR Placement history: 77% 0 or 1 prior placement; 65% long term placement Maltreatment history: 93% MH Behavioural problems: NR Foster carer experience: 68% first time FC Kinship versus non-kinship: 85% NK	Mean age: 1.5 years % female: 49 Ethnicity: NR Placement history: 88% 0 or 1 prior placement; 62% long term placement Maltreatment history: 89% MH Behavioural problems: NR Foster carer experience: 63% first time FC Kinship versus non-kinship: 83% NK
Duration of study	6 × 1.5 h visits, fortnightly, over 3 months	
Reported outcomes	Self-Regulation: Diurnal Cortisol Parent Psychological Functioning: Dutch adaptation of Parenting Stress Index (NOSI-R) (Abidin 1990) Parent–Child Relationship: Emotional Availability Scales, sensitivity; structuring; nonintrusiveness; responsiveness; involvement subscales (EAS-S/ST/N/R/I) (Biringen et al. 2014)	
Reported primary aim (s) of intervention	To improve or optimise the relationship between foster carers and foster child	
Reported aim (s) of study(s)	To use a randomised controlled pre-test post-test design to compare FFI with Care as Usual	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	Computer generated list. Significant differences reported in baseline outcome variables
Allocation bias	Low	Child foster carer couples entered our study by giving them the first available number from a computer generated list, which randomly assigned the case to the FFI or CAU condition
Performance bias	High	No blinding reported. TAU control group used. No participant blinding possible
Detection bias	Low	Researcher and assessors blind to condition, range of assessors used
Attrition bias	High	Significant attrition in participants at outcome assessment (> 20%)
Reporting bias	Unclear	No pre-registration of trial protocol or outcome variables

Characteristics of IY: incredible years

Published article/s	Bywater et al. (2011)	
Methods	Unregistered RCT	
Setting	United Kingdom	
Inclusion criteria	Foster carers with children aged 2–17	
Participant characteristics	N = 46 Foster and Kinship Carers with foster children aged 2–17 (M = 9.5 years)	
Intervention characteristics	IY intervention	TAU
Group n	n = 29	n = 17
Baseline demographics	Mean age: 8.9 years	Mean age: 10.4 years
	% female: 48	% female: 47
	Ethnicity: NR	Ethnicity: NR
	Placement history: duration current placement 1.8 years	Placement history: duration current placement 2.1 years
	Maltreatment history: NR	Maltreatment history: NR
	Behavioural problems: NR	Behavioural problems: NR
	Foster carer experience: 4.8 years	Foster carer experience: 6.9 years
	Kinship versus non-kinship: NR	Kinship versus non-kinship: NR
Duration of study	12 × 2 h weekly group training sessions	
Reported outcomes	Behaviour Problems: Eyberg Child Behavior Inventory—Intensity (ECBI-I) (Eyberg and Pincus 1999); Strengths and Difficulties Questionnaire, Hyperactivity subscale (SDQ-H) (Goodman 2001)	
	Psychological Functioning: Strengths and Difficulties Questionnaire—parent report (SDQ-P) (Goodman et al. 2004)	
	Parental Skills: Parenting Scale (PS) (Arnold et al. 1993)	
	Parent Psychological Functioning: Beck Depression Inventory (Beck et al. 1996)	
Reported primary aim (s) of intervention	Strengthening ‘parenting’ skills, with the intention of preventing, reducing and/or treating conduct problems among children aged 2–8 years while increasing their social competence	
Reported aim (s) of study(s)	Assess ‘parenting’ competency, carers’ depression levels, child behaviour and service use costs	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	Randomization, at the individual level, with a 2:1 intervention to waiting-list control condition; Significant difference at baseline of outcome variable (ECBI) ‘accounted for by ANCOVA’
Allocation bias	Low	Allocated to either condition using a random number generator unless they had commitments ruling out possible attendance at a specific group (n = 6)
Performance bias	High	No blinding reported. TAU control group used. No participant blinding possible
Detection bias	High	No reported blinding of researchers or participant assessors + self-report measures
Attrition bias	High	No consort diagram provided, no attrition reported, carers attended mean 75% session overall, unclear if there were dropouts
Reporting bias	High	Trial not registered. Multiple outcomes reported: primary outcome is the only significant child outcome. Between groups not reported (ANCOVA)

Characteristics of IY – CP: incredible years co-parenting adaptation

Published article/s	Linares et al. (2006)
Method	Unregistered RCT
Setting	United States

Inclusion/exclusion criteria	Included Foster carers with children aged 3–10 with substantiated history of maltreatment and a goal of reunification. Excluded: children with documented developmental disability (e.g. autism) or history of sexual abuse and parents with mental handicap.	
Participant characteristics	$N = 64$ Foster and Kinship Carers with foster children aged 3–10 ( $M = 6.2$ years; $SD =$ ); $N = 64$ biological parents	
Intervention characteristics	IT-CP-B intervention	TAU
Group $n$	$n = 40$	$n = 24$
Baseline demographics ( $N$ )	Mean age: 8.9 years % female: 48% Ethnicity: NR Placement history: NR Maltreatment history: 83% neglected, 6% abused Maltreatment history ( $n$ ): % neglected 71, abused 29%      Maltreatment history ( $n$ ): % neglected 100, abused 0% Behavioural problems: NR Foster carer experience: NR Kinship versus non-kinship: 100% non-kinship	
Duration of study	12 × 2 h weekly group training sessions + 12 1 h weekly co-parenting sessions	
Reported outcomes	Behaviour Problems: Eyberg Child Behavior Inventory—Intensity (ECBI), Sutter-Eyberg Student Behavior Inventory—Revised (SESBI-R) (Eyberg and Pincus 1999); Child Behaviour Checklist—externalising subscale (CBCL-E) (Achenbach and Rescorla 2001) Co-parenting relationship: composite measure (CoP) (Composite measure of combination of items from the FFSS, FACES III, and 2 newly developed items) (Linares et al. 2006) Parental Skills: Parenting Practices Interview (PPI) (Webster-Stratton 1998)	
Reported primary aim (s) of intervention	Treat externalizing problems; Enhance service integration by adopting a joint training format (biological and foster parent pairs) to meet the special composition of families with children in foster placement; increase positive parenting practices, increase collaborative parenting practices	
Reported aim (s) of study(s)	Evaluate the effectiveness of a child-focused adaptation of the Incredible Years Child Training program to reduce physical aggression	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Unclear	No detailed of reporting on randomisation methods. Significant differences in baseline demographics of groups that may favour intervention
Allocation bias	Unclear	Allocation sequence not described in detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU control group used. No participant blinding possible
Detection bias	Unclear	Blinded assessment teams reported, no blinding of parent assessment, no clinician assessed measures
Attrition bias	Low	Consort diagram provided. < 20% post intervention attrition, > 20% at follow up, reasons provided for drop outs. No demographic comparison of drop outs. ITT reported
Reporting bias	Unclear	No pre-registration of trial protocol or outcome variables

#### Characteristics of KEEP1: Keeping Foster Parents Trained and Supported

Published article/s	Chamberlain et al. (1992)
Method	Unregistered RCT
Setting	United States
Inclusion/exclusion criteria	Foster Carers and children with > 3 month placement,
Participant characteristics	$N = 72$ Foster and Kinship Carers with foster children aged 4–18 ( $M = 10.8$ years)

Intervention characteristics	KEEP	TAU + Monthly additional stipend	TAU
Group <i>n</i>	<i>n</i> = 31	<i>n</i> = 14	<i>n</i> = 27
Baseline demographics ( <i>N</i> )	Mean age: 10.8 years (range = 4–18) % female: 61 Ethnicity: Cauc: 86%; Af Am 6%, Hisp 4%, Nat Am 4% Placement history: Number of prior placements: <i>M</i> = 1.5, <i>SD</i> = 17 (range 0–9) Maltreatment history: reason for OOHC: neglect 33%, Abuse 18%; Sexual Abuse 16% Conduct problems: Runaway 22%, Drug/alcohol 6%, Felony 5%, Sexually abusive 18%, Dangerous 22% Behavior Problems: 41%, expelled 97% Psychological Problems: Suicide attempt 5%, Danger to self 11%, Danger to others: mild 18%, severe 3% Foster carer experience: 7 years, number of previous children cared for <i>M</i> = 21 (range 1–215) Kinship versus non-kinship: NR		
Duration of study	2 h weekly group training sessions for an average of 9 months		
Reported outcomes	Behaviour Problems: Parent Daily Report Checklist: (PDR) (Chamberlain and Reid 1987) Placement: Foster carer dropout rate; Number of successful days in foster care		
Reported primary aim (s) of intervention	Not stated		
Reported aim (s) of study(s)	An experimental test to measure the impact on 72 children of a \$70 per month supplement... and of increased training and support of foster parents		

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	High	No detailed reporting on randomisation. No sig differences in demographics between groups. Large and significant difference in baseline behaviour problems. No reporting on randomisation or allocation methods
Allocation bias	Unclear	Allocation sequence not described in detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU control. No participant blinding possible
Detection bias	High	No reported blinding telephone interviewers + non-blinding of participant assessors (TAU)
Attrition bias	High	Attrition rates not reported for Behaviour Outcome
Reporting bias	Unclear	Trial not registered

Characteristics of KEEP2: keeping foster parents trained and supported

Published article/s	Chamberlain et al. (2008a, b) and Price et al. (2008)	
Method	Unregistered RCT	
Setting	United States, Community group setting	
Inclusion criteria	Foster carers with children aged 5–12 years old, in out of home care for > 30 days, with 1 or more other children in home. Exclusion: Not severely mentally or physical handicapped, did not previously receive KEEP intervention	
Participant characteristics	<i>N</i> = 700 Foster and Kinship Carers with foster children aged 4–13 ( <i>M</i> = 8.8 years)	
Intervention characteristics	KEEP2	TAU
Group <i>n</i>	<i>n</i> = 359	<i>n</i> = 351
Baseline demographics	Mean age: 8.9 years % female: 50 Ethnicity: Cauc: 20%; Af Am 23%, Hisp 35%, Nat Am 1% Placement history: No. prior: <i>M</i> = 2.95 Maltreatment history: NR Behavioural problems: NR Foster carer experience: 7 years Kinship versus non-kinship: 32% Kin	Mean age: 8.7 years % female: 54 Ethnicity: Cauc: 25%; Af Am 19%, Hisp 30%, Nat Am 1% Placement history: No. prior: <i>M</i> = 2.82. Maltreatment history: NR Behavioural problems: NR Foster carer experience: 7 years Kinship versus non-kinship: 36% Kin
Duration of study	2 h weekly group training sessions for (16 weeks)	

Reported outcomes	Behaviour Problems: Parent Daily Report Checklist: (PDR) (Chamberlain and Reid 1987) Parenting Skills: ratio of positive reinforcement to discipline (PROP-REINF) (Chamberlain and Reid 1987); Placement: Number of positive placement exits; number of negative placement exits
Reported primary aim (s) of intervention	A culturally relevant universal intervention offered to foster parents receiving a new foster care placement designed to improve parenting practices, reduce behaviour problems and reduce placement disruption
Reported aim (s) of study(s)	Examine the impact of a foster parent training and support on rate of positive reinforcement and behaviour problems. Examine mediating role of positive reinforcement on behaviour problems and mediating role of baseline behaviour on intervention effects. Examine intervention impact on placement changes and placement disruption risks associated with children's placement histories

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Unclear	No detailed reporting on randomisation. No sig differences in demographics or outcome measures at baseline
Allocation bias	Unclear	Allocation sequence not described in detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU control. No participant blinding possible
Detection bias	High	No reported blinding telephone interviewers + non-blinding of participant assessors (TAU)
Attrition bias	Low	Attrition rate (< 20%). Comparable attrition rates in treatment and control. ITT analysis reported
Reporting bias	Unclear	Trial is registered. Primary Outcome measure (PDR) is same in trial registry and report

#### Characteristics of KEEP3: keeping foster parents trained and supported

Published article	Price et al. (2015)	
Method	Registered RCT; clinical trials.gov as 'Community Implementation of KEEP: Fidelity and Generalization of Parenting (KEEP)'	
Setting	United States, Community group setting	
Inclusion criteria	Foster carers with children aged 5–12 years old, in out of home care for > 30 days, with 1 or more other children in home. Exclusion: Not severely mentally or physical handicapped, did not previously receive KEEP intervention	
Participant characteristics	N = 354 Foster and Kinship Carers with foster children aged 5–12 (M = 7.55)	
Intervention characteristics	KEEP-3 intervention	TAU
Group n	179	175
Baseline demographics	Mean age: 7.8 years (SD = 2.5) % female: 47 Ethnicity: Cau: 11%; Af Am 23%, Hisp 46%, Nat Am 1%	Mean age: 7.3 years (SD = 2.3) % female: 49 Ethnicity: Cau: 18%; Af Am 12%, Hisp 51%, Nat Am 1%
	Placement history: No. prior: NR Maltreatment history: NR Behavioural problems: NR Foster carer experience: 3.1 years Kinship versus non-kinship: 45% kin	Placement history: No. prior: NR Maltreatment history: NR Behavioural problems: NR Foster carer experience: 3.8 years Kinship versus non-kinship: 50% Kin
Duration of study	16 × 1.5 h weekly group training sessions	
Registered outcomes	Behaviour Problems: Parent Daily Report Checklist: (PDR) (Chamberlain and Reid 1987); Parent Psychological Functioning: Parent Daily Report Checklist, parental stress subscale (PDR-S) (Chamberlain and Reid 1987)	
Reported outcomes	Behaviour Problems: Parent Daily Report Checklist: (PDR: Chamberlain and Reid 1987); Parent Psychological Functioning: Parent Daily Report Checklist, parental stress subscale (PDR-S) (Chamberlain and Reid 1987)	
Registered primary aim(s) of study	(1) Do the effects of KEEP generalize (concurrently) to other children currently in foster and kin intervention homes and lead to reductions in overall levels of behaviour problems? (2) Does KEEP continue to have effects after the completion of the intervention and generalize (temporally) to new children who enter the homes of these families at a later point in time? (3) Can KEEP be delivered and maintained in a manner that preserves the goals and quality of the intervention when implemented by a community agency in a real-world system of care	



Reported aim(s) of study(s)	(1) examine the effectiveness of KEEP in reducing child behaviour problems, as delivered by a community agency; (2) determine if the effects of KEEP generalize to more than one child in the same home; (3) examine the effectiveness of KEEP in reducing parental stress associated with child behaviour problems	
Reported primary aim(s) of intervention	Increasing the use of positive reinforcement, consistent use of non-harsh discipline methods, monitoring and control of peer associations	
Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Unclear	If caregivers agreed to participate, they were asked to sign a consent form and to select an envelope from among several unidentified and sealed envelopes. Each envelope contained a sheet of paper on which was either the word “intervention” or the word “control
Allocation bias	Low	Allocation described in sufficient detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU control. No participant blinding possible
Detection bias	Unclear	Reported blinding telephone interviewers (trial) but non-blinding of participant assessors (TAU)
Attrition bias	Low	Attrition < 20%). Comparable attrition rates. ITT analysis reported
Reporting bias	High	Primary OM is same in trial registry and report (PDR)

#### Characteristics of KITS: kids in transition to school

Published articles	Pears et al. (2012, 2013, 2014, 2016)	
Method	Registered RCT; clinical trials.gov as ‘KITS: School Readiness in Foster Care Efficacy Trial (KITS). ID: NCT00688129	
Setting	United States, Community group setting	
Inclusion criteria	Foster carers and foster children in kinship or non-kinship foster care entering kindergarten, not involved in treatment associated with KITS	
Participant characteristics	N = 192 Foster Carers with foster children aged 4–6 (M = 5.3 years)	
Intervention characteristics	KITS intervention	TAU
Group n	n = 102	n = 90
Baseline demographics	Mean age: 5.3 years % female: 46 Ethnicity: Cau: 55%; Af Am 1%, Hisp 30%, Nat Am 2% Placement history: No. prior: NR Maltreatment history: NR Behavioural problems: NR Foster carer experience: 3.1 years Kinship versus non-kinship: 38% kin	Mean age: 5.3 years % female: 52 Ethnicity: Cau: 51%; Af Am 0%, Hisp 31%, Nat Am 0% Placement history: No. prior: NR Maltreatment history: NR Behavioural problems: NR Foster carer experience: 3.1 years Kinship versus non-kinship: 39% kin
Duration of study	Foster children: 16 × 2 h twice-weekly group sessions + 8 × 1 h weekly session (24 sessions, 16 weeks); Foster carers: 8 × 2 h session bi-weekly (8 sessions, 16 weeks)	
Registered outcomes	Primary: (1) Early literacy (2) Social skills (3) Regulatory Skills (4) Caregiver involvement Secondary: (1) Academic achievement (2) Peer relations	
Reported outcomes	Behaviour problems: Child Behavior Checklist—Oppositional and Aggressive subscales (CBCL-O + A); Teacher Report Form—Oppositional and Aggressive subscales (TRF—O + A) (Achenbach and Rescorla 2001); Conners Teacher Rating Scale—Revised, Oppositional subscale (CTRS-S—O) Academic Functioning: Early literacy skills, Concepts About Print Assessment (CAPS) (Clay 2000); Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Good et al. 2003); Caregiver rating of literacy skills. (CR) Social Competence: Preschool Penn Interactive Peer Play Scale (Fantuzzo and Perlman 2007); Child Behavior Checklist—Social Competence (CBCL-SC) (Achenbach and Rescorla 2001); Emotion Understanding (Fries and Pollak 2004) Psychological Functioning: Self-perception profile for children, global self-worth subscale (SPCC) (Harter 1985, 1988) Self-regulation: Composite of three subcomponents: Emotion regulation (composite measure); Behaviour Regulation (composite measure), Inhibitory control (composite measure) Delinquent Behaviour: Positive attitudes towards alcohol use in third grade (Pears et al. 2016); Positive attitudes towards antisocial behaviour in third grade (Pears et al. 2016); Involvement with deviant peers at third grade (Pears et al. 2016)	

Registered primary aim (s) of intervention	(1) Early literacy skills (Time Frame: pre-intervention, pre-kindergarten, end of kindergarten); (2) Social skills (Time Frame: pre-intervention, pre-kindergarten, end of kindergarten); (3) Regulatory skills (Time Frame: pre-intervention, pre-kindergarten, end of kindergarten); (4) Caregiver involvement (Time Frame: pre-intervention, pre-kindergarten, end of kindergarten)
Reported primary aim (s) of intervention	Improving children's early literacy, prosocial, and self-regulation skills, as well as caregiver involvement and parenting skills
Reported aim (s) of study(s).	A randomized efficacy trial of a school readiness intervention... designed to increase literacy, social and self-regulation in children before kindergarten entry

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Unclear	No details of reporting on randomisation. Baseline characteristics of both groups reported, with no significant differences
Allocation bias	Unclear	Allocation sequence not described in detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU non-active control. No participant blinding possible
Detection bias	Low	Researchers reported as blind to condition + non-blinding of participant assessors, multiple outcome assessors (child, researcher, parent, teacher) (TAU)
Attrition bias	Unclear	Limited reporting on attrition or drop-out rate and ITT analysis
Reporting bias	Unclear	Trial registry did not state specific outcome instruments used, only domains. Some pre-registered outcomes not reported, Risk of inflation bias

#### Characteristics of MSS: middle school success

Published articles	Kim et al. (2013), Kim and Leve (2011) and Smith et al. (2011)	
Method	Registered RCT; clinical trials.gov as "Prevention Program for Problem Behaviors in Girls in Foster Care" ID: NCT00239837	
Setting	United States, Community group setting	
Inclusion criteria	Foster carers, kinship carers and foster children	
Participant characteristics	$N = 100$ Foster Carers and foster children aged 10–12 in final year of junior school ( $M = 11.54$ years; $SD = 0.48$ )	
Intervention characteristics	MSS intervention	TAU
Group $n$	$n = 48$	$n = 52$
Baseline demographics	Mean age ( $n$ ): 11.59 years % female ( $n$ ): 100 Behavioural problems ( $n$ ): EXT 16.06 Ethnicity ( $N$ ): Cau: 63%; Af Am 4%, Hisp 10%, Nat Am 4% (no sig. diff between groups) Placement history ( $N$ ): Time in foster care: $M = 2.9$ years ( $SD = 2.24$ ). No. of placements $M = 1.37$ ( $SD = 0.66$ ) (no sig. difference between groups) Maltreatment history ( $N$ ): 50% self-reported physical/sexual abuse, > 90% self-reported neglect, 32% reported physical and sexual abuse and neglect. Neglect sig. higher in control condition $t(95) = -2.65$ . (No sig. differences for Physical/Sexual abuse) Foster carer experience ( $N$ ): 3.1 years (no sig. diff between groups) Kinship versus non-kinship ( $N$ ): 28% (no sig. diff between groups)	Mean age ( $n$ ): 11.48 years % female ( $n$ ): 100 Behavioural problems ( $n$ ): EXT 14.38
Duration of study	Foster children: $6 \times 2$ h twice-weekly group sessions caregiver, $6 \times 2$ h twice-weekly skill building sessions for foster children. 20 2-hour follow intervention meetings for caregivers, 20 2-h follow intervention meetings for caregivers	
Registered outcomes	Substance use (child self-report); Delinquency Self-Report Delinquency Scale (SRD) (Roweton 1990); Placement Change	

Reported outcomes	Behaviour Problems: Child Behavior Checklist, average of Externalising and Internalising subscales (CBCL E + I) (Achenbach and Rescorla 2001) (Parent Daily report—externalising subscale (PDR—E) (Chamberlain and Reid 1987) Psychological Functioning: Parent Daily report—internalising subscale (PDR—I) (Chamberlain and Reid 1987) Social Competence: Parent Daily Report Checklist—prosocial subscale (PDR—PS) (Chamberlain and Reid 1987) Placement: Number of placement changes (6–12 months) Delinquent Behaviours: Self-report Delinquency Scale (SRD) (Roweton 1990); Substance Use, child self-report (Tobacco, Marijuana, Alcohol); Substance use: (Tobacco and Marijuana use, child self-report); Health-Risking Sexual Behaviours (8 item questionnaire in child interview)
Registered primary aim (s) of intervention	The intervention targets include preventing delinquency, initiation of substance use, participation in risky sexual behavior, school truancy and failure, and mental health problems
Reported primary aim (s) of intervention	Targeted the prevention of internalizing and externalizing problems during the transition to middle school to help prevent more serious, longer term outcomes such as delinquency, substance use, and high-risk sexual behavior in later middle school
Reported aim (s) of study(s).	(a) (hypothesize) fewer internalizing problems, fewer externalizing problems, and more prosocial behavior at 6-months post-baseline compared to the control girls (Smith et al. 2011). (b) (evaluate) efficacy of the Middle School Success intervention (MSS) for reducing substance use and delinquency among girls in foster care, using a randomized controlled trial (RCT) design (Kim and Leve 2011). (c) (examine) effects of (MSS) intervention, on foster-girls health-risking sexual behavior, using a randomized controlled trial (RCT) design (Kim et al. 2013)

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	A coin flip was used for the randomization process in lieu of other randomization processes (e.g. computerized) to allow for a community member to be present during the process and to provide transparency and ease of understanding to caseworkers and study participants. No significant differences at baseline
Allocation bias	Low	Randomisation was undertaken transparently
Performance bias	High	No blinding reported. TAU control. No participant blinding possible
Detection bias	Low	Interviews: ‘multimethod, multiagent assessment approach that consisted of a standardized interview and questionnaires for... girl... foster parent... girl’s caseworker... and collection of child welfare records... assessors who were blind to the participants’ group assignments’
Attrition bias	Low	15/98 dropouts, (< 20%). Equivalent of ITT analysis used
Reporting bias	High	Trial was registered, but several outcome measures reported were not registered, and some registered outcomes were not reported

Characteristics of PCIT: parent child interaction therapy

Published articles	Mersky et al. (2015, 2016)		
Method	Unregistered RCT		
Setting	United States, Community group dyads		
Inclusion criteria	Foster carers with foster children 3–6 years, with externalising problems within the clinical range on ECBI		
Participant characteristics	N = 129 Foster Carers and foster children (M = 4.6 years)		
Design	G1 Brief PCIT	G2 Extended PCIT	G3 TAU
Group n	n = 48	n = 35	n = 46
Baseline demographics	Mean age (n): 4.6 years % female: NR Ethnicity(n): African Am 61% Placement history: NR Maltreatment history: NR Behaviour problems (n): In clinical range (ECBI) Foster carer experience: median = 24 months Kinship versus non-kinship: NR Adopted: NR	Mean age (n): 4.4 years % female: 56% Ethnicity (n) African Am 58% Placement history: NR Maltreatment history: NR Behaviour problems (n): In clinical range (ECBI) Foster carer experience: median = 16 months Kinship versus non-kinship: NR Adopted: NR	Mean age (n): 4.6 years % female: 63% Ethnicity(n): African Am 63% Placement history: NR Maltreatment history: NR Behaviour problems (N): In clinical range (ECBI) Foster carer experience: median = 37 months Kinship versus non-kinship: NR Adopted: NR

Duration of study	Foster carers and Foster Children: 2 × 7 h training, 4 × 20 min weekly phone sessions + 4 bi-weekly 20 min phone sessions, G2 also had additional 1 × 7 h training and 6 × bi-weekly additional phone sessions.
Reported outcomes	Behaviour Problems: Child Externalising Behaviour Problems (CBCL-E) (Achenbach and Rescorla 2001); Eyberg Child Behavior Inventory, intensity subscale (ECBI-I); Eyberg Child Behavior Inventory, problem subscale (ECBI-P) (Eyberg and Ross 1978) Psychological Functioning: Child Behavior Checklist—Internalising subscale (CBCL-I) (Achenbach and Rescorla 2001) Parent–Child Relationship: Dyadic Parent–Child Interaction Coding System, labelled praise, negative talk, positive composite, negative composite (DPICSII—LP/NY/Pos/Neg) (Eyberg et al. 2013) Parent psychological functioning: Parenting Stress Index, total stress scale; parental distress, parent–child dysfunctional interaction, difficult child (PSI/PD/DI/DC) (Abidin 1990)
Reported primary aim(s) of intervention	Help caregivers modify their attitudes towards parenting and develop behaviour management skills that alter parent–child exchanges and shape child behaviors
Reported aim (s) of study(s)	“...extend research on PCIT by examining intervention effects on internalizing symptoms” ‘hep foster parents access empirically supported behavior management training ...foster children access empirically supported mental health services’

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	Low covariate randomization procedure used. Baseline characteristics of both groups reported, with no significant differences except for two variables controlled for in analysis
Allocation bias	Unclear	Allocation sequence not described in detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU control. No participant blinding possible
Detection bias	Low	Research assessors reported blinded + clinician and self-report assessment used
Attrition bias	Low	Attrition < 20%. ITT analysis performed; well reported, non-sig. differences in groups at baseline
Reporting bias	Unclear	No pre-registration of trial protocol or outcome variables, unclear reporting of between group linear trends

#### Characteristics of PFR: Promoting First Relationships

Published articles	Spieker et al. (2012, 2014)	
Method	Registered RCT; registered at NIH, clinicaltrials.gov. trials ID: NCT00339365	
Setting	United States, Home Visits	
Inclusion criteria	Foster, kinship and birth parents who spoke English caring for toddlers 10–24 months, following court-ordered placement resulting in change of primary caregiver in prior 7 weeks	
Participant characteristics	<i>N</i> = 210 Foster Carers and foster children 10–24 months ( <i>M</i> = 1.5 years)	
Design	PFI Intervention	EES
Group <i>n</i>	105 allocated to PFI interventions	105 allocated to EES active control
Baseline demographics	Mean age: 1.5 years	Mean age: 1.5 years
	% female: 60	% female: 52
Duration of study	Ethnicity: Af Am 16%, Caucasian, 48% native Am 8%	Ethnicity: Af Am 13%, Caucasian 61%, native Am 6%
	Placement history: age at first removal: 10.7 m months; mean caregiver disruptions: 2.7	Placement history: age at first removal: 10.9 m months; mean caregiver disruptions: 2.7
Registered outcomes	Maltreatment history: NR	Maltreatment history: NR
	Behaviour problems:	Behaviour problems:
	Foster carer experience: <i>M</i> = 2 years	Foster carer experience: <i>M</i> = 2 years
	Kin/non-kinship/birth parent: 44/29/27%	Kin/non-kin/birth parent: 41/33/25%
Duration of study	PFR: 10 × 1–1.25 h weekly visits/sessions	EES: 3 × 1.5 h monthly visits
Registered outcomes	Primary: Toddler Attachment Sort-45 Nursing Child Assessment Satellite Training (NCAS); Teaching Scale Indicator of Parent–Child Interaction (IPCI); Brief Infant–Toddler Social Emotional Assessment; Bayley Scales of Infant Development. Secondary: Parenting Stress Index/Short Form; Patient Health Questionnaire (PHQ-9) (All measures at: baseline, 1 and 6 months post-treatment)	

Reported outcomes	<p>Attachment Behaviour: Toddler Attachment Sort-45 (TAS45) (Spieker et al. 2011)</p> <p>Behaviour Problems: Child Behavior Checklist—Externalising subscale (CBCL-E) (Achenbach and Rescorla 2001); Child Behavior Checklist—Sleep problem subscale (CBCL-SL) (Achenbach and Rescorla 2001); Brief Infant–Toddler Social and Emotional Assessment—problem behaviour scale (BITSEA-P) (Briggs-gowan et al. 2002)</p> <p>Psychological Functioning: Child Behavior Checklist—Internalising subscale (CBCL-I) (Achenbach and Rescorla 2001); Child Behavior Checklist—Other problems scale (CBCL-OT) (Achenbach and Rescorla 2001)</p> <p>Self-Regulation: Bayley Behavior Rating Scales, subscale: Emotion regulation (BRS-ER) (Bayley 1993)</p> <p>Social Competence: Brief Infant–Toddler Social and Emotional Assessment—social competence scale (BITSEA-SC) (Briggs-gowan et al. 2002); Bayley Behavior Rating Scales, subscale: Orientation (BRS-O) (Bayley 1993)</p> <p>Parent–Child Relationship: Indicator of Parent–Child Interaction, child engagement (IPCI-E); parent support (IPCI-PS) (Baggett et al. 2003)</p> <p>Parental Sensitivity: Nursing Child Assessment Teaching Scale (NCATS) (Barnard 1978) (Barnard 1978)</p> <p>Parental Knowledge: Raising a Baby (RAB) (Kelly and Korfmacher 2008)</p> <p>Parental Commitment: This Is My Baby (TIMB) (Bernard and Dozier 2011) (Bernard and Dozier 2011)</p> <p>Parent Psychological Functioning: Parent Stress Index, difficult child scale (PSI) (Abidin 1995); Parent Stress Index, parent–child dysfunction scale (PSI) (Abidin 1995)</p> <p>Placement Outcomes: Placement Stability; placement permanency</p>
Registered primary aim(s) of intervention	...promoting infant well-being, preventing emotional and behavioral problems, countering developmental delay, and reducing placement instability in young foster care children
Reported primary aim(s) of intervention	PFR helps caregivers identify possible “miscues” empathize with the child’s underlying distress, and reframe or understand the child’s behavior as reflecting an unmet need. This better understanding of the child’s cues should, in turn, lead to more responsive, nurturing care
Reported aim (s) of study(s)	This study reports on child welfare outcomes of a community based, randomized control trial of Promoting First Relationships®... 10-week relationship-based home visiting program, on stability of children’s placements and permanency status 2 years after enrollment into the study

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	Comparison condition was computer generated, blocked by caregiver type. Some differences in placement history, but not in favour of intervention
Allocation bias	Low	Randomization into the PFR intervention and the Early Education Support (EES) comparison condition was computer generated, blocked by caregiver type
Performance bias	Low	Single blinded subjects, with active control
Detection bias	Low	Research visitors were blind to condition status + self-report participant raters + clinician raters, active comparator
Attrition bias	Low	< 20% attrition for immediate outcomes. HIGH? Follow up at 6 months report higher real attrition rate
Reporting bias	Low	Pre-registered, with multiple outcome variables, some (but not all) of which were which were changed after study was conducted. However, primary registered outcome (sensitivity) was reported as primary outcome in study

Characteristics of PMTO: parent management training oregon

Published article	Maaskant et al. (2016)
Method	Unregistered RCT
Setting	The Netherlands
Inclusion criteria	Foster and kinship carers with foster children aged 4–12 currently in long-term placement at high risk of breakdown, with significant behaviour problems indicated by SDQ scores above clinical cut-off
Participant characteristics	N = 88 Foster Carers and foster children 4–12 years (M = 7.7 years)

Allocation	PMTO intervention	CAU
Group <i>n</i>	<i>N</i> = 47	<i>N</i> = 41
Baseline demographics	Mean age: 7.9 years % female: 64 Ethnicity: 'non Dutch' 39% Placement history: No. of previous placements <i>M</i> = 0.96, age at entering first placement <i>M</i> = 3.46; Duration current placement <i>M</i> = 4.4 years Maltreatment history: NR Behaviour problems: CBCL EXT T = 66.43 Foster carer experience: <i>M</i> = 7.8 years Kin/non-kinship: 17/83%	Mean age: 7.5 years % female: 50 Ethnicity: 'non Dutch' 20% Placement history: No. of previous placements <i>M</i> = 0.96, age at entering first placement <i>M</i> = 3.6, Duration current placement <i>M</i> = 3.3 years Maltreatment history: NR Behaviour problems: CBCL EXT T = 67.13 Foster carer experience: <i>M</i> = 7.2 years Kin/non kinship/birth carer: 15/85%
Duration of study	Weekly 1.5 h sessions, <i>M</i> = 21 sessions	
Reported outcomes	Behaviour Problems: Child Behavior Checklist—Total problems (CBCL) and Externalising subscale (CBCL-E) (Achenbach and Rescorla 2001); Teacher Report Form—Total Problems subscale (TRF); Externalising subscale (TRF-E) (Achenbach and Rescorla 2001) Psychological Functioning: Child Behavior Checklist—Internalising subscale (CBCL-I) (Achenbach and Rescorla 2001); Teacher Report Form, Internalising problems (TRF-I) (Achenbach and Rescorla 2001) Parent Psychological Functioning: Dutch adaptation of Parenting Stress Index (NOSI-R) (Abidin 1990) Parental Skills: Parenting Behavior Questionnaire; warmth; responsiveness; explaining; autonomy granting; strictness; discipline subscales (PBQ-W/R/E/A/S/D) (McWayne et al. 2008)	
Reported primary aim(s) of intervention	The aim of PMTO, a relative long and intensive (6–9 months, with weekly sessions) parent management training is to reduce children's problem behavior through improvement of parenting practices	
Reported aim (s) of study(s)	Investigate effectiveness of PMTO to reduce foster parenting stress. It was expected that PMTO would reduce parenting stress in foster parents, improve the quality of parenting practices, and reduce children's problem behaviour	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	Low	Randomization was undertaken by coin tossing by the second author, who was not involved in the enrolment process and blind for personal information of the eligible participants
Allocation bias	Low	All staff and counsellors of foster care organizations were blind for the randomization process. The researchers were not involved with the implementation and execution of PMTO
Performance bias	High	No blinding reported of participants. No participant blinding possible (TAU control) Foster agencies blind to allocation
Detection bias	Low	The researchers were not involved with the implementation and execution of PMTO + self-report + TAU
Attrition bias	Low	Attrition > 20%, reasons provided, ITT analysis (LOC) used
Reporting bias	High	No pre-registration of trial protocol or outcome variables. PMTO not usually associated with stated primary aim or reported primary outcome (PSI), which was only significant effect of intervention

#### Characteristics of PSB: promoting sibling bonds

Published articles	Linares et al. (2015)
Method	Unregistered RCT
Setting	USA, community setting
Inclusion criteria	Sibling pairs with same foster placement, age 5–11 years, with history of neglect with or without physical abuse



Participant characteristics	N = 22 Foster carers and foster sibling pairs 5–11 years (M = 8.15)	
Allocation	PSB intervention	TAU
Group n	n = 13	n = 9
Baseline Demographics	Mean age: 7.2/9.6 years % female (N): 26% MM, 37% FF, 37% MF Ethnicity: Af Am 37%, Hispanic 29%, Other 35% Placement history (N): age at first placement: M = 4.9, months spent in foster care M = 32 months, current placement (n): < 12 months 53.8% 12–24:30.8%, > 24 15% Maltreatment history: Neglect 100%, abuse 5% Behaviour problems: CBCL EXT percentile M = 65.4 Foster carer experience: M = 2 years Kin/non-kinship/birth carer: 37/63%	Mean age: 7.3/8.5 years % female (N): 26% MM, 37% FF, 37% MF Ethnicity: Af Am 62%, Hispanic 20%, Other 18% Placement history (N): age at first placement: M = 4.7, months spent in foster care M = 32 months current placement (n): < 12 months 44.4%, 12–24 months:11.1%, > 24 months 44.4% Maltreatment history: Neglect 78%, Abuse 22% Behaviour problems: CBCL EXT percentile M = 50 Foster carer experience: M = 2 years Kin/non-kinship/birth carer: 42/58%
Duration of study	8 × 90 min weekly sessions with foster siblings, 8 × 90 min weekly sessions with foster carers	
Reported primary aim(s) of intervention	Increasing sibling positive interaction, reducing conflict during play, and promoting conflict mediation strategies	
Reported outcomes	Behaviour problems: Sibling Aggression Scale, verbal subscale; physical subscale (SAS-V/P) (Linares 2008) Social Competence: Sibling Interaction Quality, positive interactions subscale; negative interaction subscale; conflict subscale (SIQ-P/N/C) (Kramer and Houston 1999) (Kramer 2010) Parental skills: Parent Conflict Mediation Conflict Checklist (CC) (Smith and Ross 2007)	
Reported aim (s) of study(s)	Evaluate the feasibility and short-term effectiveness of a transtheoretical intervention model targeting sibling pairs and their foster parent that integrates family systems, social learning theory, and a conflict mediation perspective	

Item	Judgement	Comment
<i>Risk of bias</i>		
Randomisation bias	High	No details of reporting on randomisation. Significant differences in child maltreatment history and medication that may favour intervention (due to small groups?)
Allocation bias	Unclear	Allocation sequence not described in detail. Described as randomly allocated
Performance bias	High	No blinding reported. TAU control. No participant blinding possible
Detection bias	Low	Assessor blinding + clinician rated outcomes. + self-report + TAU
Attrition bias	Low	Attrition 18–22%, ITT analysis used
Reporting bias	Unclear	No pre-registration of trial protocol or outcome variables

## References

- Abidin, R. R. (1990). *Parenting stress index-short form*. Charlottesville, VA: Pediatric Psychology Press.
- Abidin, R. R. (1995). *The Parenting Stress Index*. Odessa, FL: Psychological Assessment Resources.
- Achenbach, T. M., & Rescorla, L. (2001). *ASEBA school-age forms and profiles*. Burlington: ASEBA.
- Appleyard, K., Egeland, B., Dulmen, M. H., & Alan Sroufe, L. (2005). When more is not better: The role of cumulative risk in child behavior outcomes. *Journal of Child Psychology and Psychiatry*, 46(3), 235–245.
- Arnold, D. S., O’Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, 5(2), 137–144. <https://doi.org/10.1037/1040-3590.5.2.137>.
- Baggett, K., Carta, J., & Horn, E. (2003). *Indicator of parent–child interaction (IPCI)*. Lawrence, KS: University of Kansas.
- Barnard, K. (1978). *Nursing child assessment teaching scale*. Seattle: University of Washington School of Nursing.
- Bayley, N. (1993). *Bayley scales of infant development: Manual*. San Antonio: Psychological Corporation.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory-II (BDI-II)*. San Antonio, TX: Psychological Corporation.
- Begley, C. G., & Ioannidis, J. P. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. *Circulation Research*, 116(1), 116–126. <https://doi.org/10.1161/circresaha.114.303819>.



- Bernard, K., & Dozier, M. (2011). This is my baby: Foster parents' feelings of commitment and displays of delight. *Infant Mental Health Journal, 32*(2), 251–262.
- Bick, J., & Dozier, M. (2013). The effectiveness of an attachment-based intervention in promoting foster mothers' sensitivity toward foster infants. *Infant Mental Health Journal, 34*(2), 95–103. <https://doi.org/10.1002/imhj.21373>.
- Biringen, Z., Derscheid, D., Vliegen, N., Closson, L., & Easterbrooks, M. A. (2014). Emotional availability (EA): Theoretical background, empirical research using the EA Scales, and clinical applications. *Developmental Review, 34*(2), 114–167. <https://doi.org/10.1016/j.dr.2014.01.002>.
- Bishop, D. V. M., & Thompson, P. A. (2016). Problems in using p-curve analysis and text-mining to detect rate of p-hacking and evidential value. *PeerJ, 4*, e1715. <https://doi.org/10.7717/peerj.1715>.
- Blythe, S. L., Wilkes, L., & Halcomb, E. J. (2014). The foster carer's experience: An integrative review. *Collegian, 21*(1), 21–32. <https://doi.org/10.1016/j.colegn.2012.12.001>.
- Briggs-gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2002). *Brief Infant-Toddler Social and Emotional Assessment (BITSEA) manual, version 2.0*.
- Briskman, J., Castle, J., Blackeby, K., Bengo, C., Slack, K., Stebbens, C., et al. (2012). *Randomised controlled trial of the fostering changes programme*. London: Department for Education. [www.gov.uk/government/publications/randomised-controlled-trial-of-the-fostering-changes-programme](http://www.gov.uk/government/publications/randomised-controlled-trial-of-the-fostering-changes-programme).
- Buehler, C., Cox, M. E., & Cuddeback, G. (2003). Foster parents' perceptions of factors that promote or inhibit successful fostering. *Qualitative Social Work, 2*(1), 61–83. <https://doi.org/10.1177/1473325003002001281>.
- Bywater, T., Hutchings, J., Linck, P., Whitaker, C., Daley, D., Yeo, S. T., et al. (2011). Incredible years parent training support for foster carers in Wales: A multi-centre feasibility study (Multi-center Study; Randomized Controlled Trial; Research Support, Non-U.S. Gov't). *Child: Care, Health and Development, 37*(2), 233–243. <https://doi.org/10.1111/j.1365-2214.2010.01155.x>.
- Chamberlain, P., Moreland, S., & Reid, K. (1992). Enhanced services and stipends for foster parents: Effects on retention rates and outcomes for children. *Child Welfare: Journal of Policy, Practice, and Program, 71*(5), 387–401.
- Chamberlain, P., Price, J., Leve, L. D., Laurent, H., Landsverk, J. A., & Reid, J. B. (2008a). Prevention of behavior problems for children in foster care: Outcomes and mediation effects. *Prevention Science, 9*(1), 17–27. <https://doi.org/10.1007/s1121-007-0080-7>.
- Chamberlain, P., Price, J., Reid, J., & Landsverk, J. (2008b). Cascading implementation of a foster and kinship parent intervention. *Child Welfare, 87*(5), 27–48.
- Chamberlain, P., & Reid, J. B. (1987). Parent observation and report of child symptoms. *Behavioral Assessment, 9*(1), 97–109.
- Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685.
- Clarke, M., & Oxman, A. (2000). *Cochrane reviewers' handbook*. Oxford: Update Software.
- Clay, M. M. (2000). *Concepts about print*. Portsmouth, NH: Heinemann Education.
- Cloitre, M., Stolbach, B. C., Herman, J. L., Kolk, B. V. D., Pynoos, R., Wang, J., et al. (2009). A developmental approach to complex PTSD: Childhood and adult cumulative trauma as predictors of symptom complexity. *Journal of Traumatic Stress, 22*(5), 399–408. <https://doi.org/10.1002/jts.20444>.
- Craven, P. A., & Lee, R. E. (2006). Therapeutic interventions for foster children: A systematic research synthesis. *Research on Social Work Practice, 16*(3), 287–304. <https://doi.org/10.1177/1049731505284863>.
- DeAngelis, C. D. M. D. M. P. H., Drazen, J. M. M. D., Frizelle, F. A. M. M. F., Haug, C. M. D. P. M., Hoey, J. M. D., Horton, R. F., et al. (2004). Clinical Trial Registration: A statement from the international committee of medical journal editors. *JAMA, 292*(11), 1363–1364.
- Dorsey, S., Farmer, E. M. Z., Barth, R. P., Greene, K. M., Reid, J., & Landsverk, J. (2008). Current status and evidence base of training for foster and treatment foster parents. *Children and Youth Services Review, 30*(12), 1403–1416. <https://doi.org/10.1016/j.childyouth.2008.04.008>.
- Dozier, M., Lindhiem, O., Lewis, E., Bick, J., Bernard, K., & Peloso, E. (2009). Effects of a foster parent training program on young children's attachment behaviors: Preliminary evidence from a randomized clinical trial. *Child and Adolescent Social Work Journal, 26*(4), 321–332. <https://doi.org/10.1007/s10560-009-0165-1>.
- Dozier, M., Peloso, E., Lewis, E., Laurenceau, J.-P., & Levine, S. (2008). Effects of an attachment-based intervention of the cortisol production of infants and toddlers in foster care. *Development and Psychopathology, 20*(3), 845–859. <https://doi.org/10.1017/S0954579408000400>.
- Dozier, M., Peloso, E., Lindhiem, O., Gordon, M. K., Manni, M., Sepulveda, S., et al. (2006). Developing evidence-based interventions for foster children: An example of a randomized clinical trial with infants and toddlers. *Journal of Social Issues, 62*(4), 767–785. <https://doi.org/10.1111/j.1540-4560.2006.00486.x>.
- Dozier, M., & Stovall, K. C. (1997). *Parent attachment diary*. Unpublished manuscript, University of Delaware.
- Everson-Hock, E. S., Jones, R., Guillaume, L., Clapton, J., Goyder, E., Chilcott, J., et al. (2012). The effectiveness of training and support for carers and other professionals on the physical and emotional health and well-being of looked-after children and young people: A systematic review. *Child: Care, Health and Development, 38*(2), 162–174. <https://doi.org/10.1111/j.1365-2214.2011.01247.x>.
- Eyberg, S. M., & Pincus, D. (1999). *Eyberg child behavior inventory and sutter-eyberg student behavior inventory-revised: Professional manual*. Psychological Assessment Resources.
- Eyberg, S. M., & Ross, A. W. (1978). Assessment of child behavior problems: The validation of a new inventory. *Journal of Clinical Child & Adolescent Psychology, 7*(2), 113–116.
- Eyberg, S. M., Nelson, M. M., Ginn, N. C., Bhuiyan, N., & Boggs, S. R. (2013). *Dyadic Parent-Child Interaction Coding System: Comprehensive Manual for Research and Training* (4th ed.). Gainesville, FL: PCIT International.
- Fantuzzo, J., & Perlman, S. (2007). The unique impact of out-of-home placement and the mediating effects of child maltreatment and homelessness on early school success. *Children and Youth Services Review, 29*(7), 941–960. <https://doi.org/10.1016/j.childyouth.2006.11.003>.
- Feise, R. J. (2002). Do multiple outcome measures require p-value adjustment? *BMC Medical Research Methodology, 2*(1), 1.
- Festinger, T., & Baker, A. J. L. (2013). The quality of evaluations of foster parent training: An empirical review. *Children and Youth Services Review, 35*(12), 2147–2153. <https://doi.org/10.1016/j.childyouth.2013.10.009>.
- Fisher, P. A., Burraston, B., & Pears, K. (2005). The early intervention foster care program: permanent placement outcomes from a randomized trial (Clinical Trial; Randomized Controlled Trial; Research Support, U.S. Gov't, P.H.S.). *Child Maltreatment, 10*(1), 61–71. <https://doi.org/10.1177/1077559504271561>.
- Fisher, P. A., Gunnar, M. R., Chamberlain, P., & Reid, J. B. (2000). Preventive intervention for maltreated preschool children: Impact on children's behavior, neuroendocrine activity, and foster parent functioning. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(11), 1356–1364. <https://doi.org/10.1097/00004583-200011000-00009>.

- Fisher, P. A., & Kim, H. K. (2007). Intervention effects on foster preschoolers' attachment-related behaviors from a randomized trial. *Prevention Science*, 8(2), 161–170. <https://doi.org/10.1007/s11121-007-0066-5>.
- Ford, J. D., & Courtois, C. A. (Eds.). (2013). *Treating complex stress disorders in children and adolescents: Scientific foundations and therapeutic models*. New York, NY: Guilford Press.
- Fraser, J. G., Lloyd, S., Murphy, R., Crowson, M., Zolotor, A. J., Coker-Schwimmer, E., et al. (2013). A comparative effectiveness review of parenting and trauma-focused interventions for children exposed to maltreatment. *Journal of Developmental and Behavioral Pediatrics*, 34(5), 353–368.
- Frick, R. W. (1996). The appropriate use of null hypothesis testing. *Psychological Methods*, 1(4), 379.
- Fries, A. B. W., & Pollak, S. D. (2004). Emotion understanding in postinstitutionalized Eastern European children. *Development and Psychopathology*, 16(2), 355–369. <https://doi.org/10.1017/S0954579404044554>.
- Gavita, O. A., David, D., Bujoreanu, S., Tiba, A., & Ionutiu, D. R. (2012). The efficacy of a short cognitive-behavioral parent program in the treatment of externalizing behavior disorders in Romanian foster care children: Building parental emotion-regulation through unconditional self- and child-acceptance strategies. *Children and Youth Services Review*, 34(7), 1290–1297. <https://doi.org/10.1016/j.chilcyouth.2012.03.001>.
- Goemans, A., van Geel, M., & Vedder, P. (2015). Over three decades of longitudinal research on the development of foster children: A meta-analysis. *Child Abuse and Neglect*, 42, 121–134. <https://doi.org/10.1016/j.chiabu.2015.02.003>.
- Goldbeck, L., & Vitiello, B. (2011). Reporting clinical trials of psychosocial interventions in child and adolescent psychiatry and mental health. *Child and Adolescent Psychiatry and Mental Health*, 5(1), 4. <https://doi.org/10.1186/1753-2000-5-4>.
- Good, R. H., Kaminski, R. A., Smith, S., Laimon, D., & Dill, S. (2003). *Dynamic indicators of basic early literacy skills*. Colorado: Sopris West Educational Services Longmont.
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(11), 1337–1345.
- Goodman, R., Ford, T., Corbin, T., & Meltzer, H. (2004). Using the Strengths and Difficulties Questionnaire (SDQ) multi-informant algorithm to screen looked-after children for psychiatric disorders. *European Child and Adolescent Psychiatry*, 13(Suppl2), II25–II31. <https://doi.org/10.1007/s00787-004-2005-3>. (validation studies).
- Greenberg, M. T., Speltz, M. L., DeKlyen, M., & Jones, K. (2001). Correlates of clinic referral for early conduct problems: Variable- and person-oriented approaches. *Development and Psychopathology*, 13(02), 255–276.
- Gutman, L. M., Sameroff, A. J., & Cole, R. (2003). Academic growth curve trajectories from 1st grade to 12th grade: Effects of multiple social risk factors and preschool child factors. *Developmental Psychology*, 39(4), 777–790. <https://doi.org/10.1037/0012-1649.39.4.777>.
- Haidich, A. B. (2010). Meta-analysis in medical research. *Hippokratia*, 14(Suppl 1), 29–37.
- Harter, S. (1985). *Manual for the self-perception profile for children*. Denver, CO: University of Denver.
- Harter, S. (1988). *Manual for the self-perception profile for adolescents*. Denver, CO: University of Denver.
- Higgins, P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., et al. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343, d5928.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med*, 2(8), e124.
- Ioannidis, J. P., Patsopoulos, N., & Rothstein, H. (2008). Research methodology: Reasons or excuses for avoiding meta-analysis in forest plots. *BMJ: British Medical Journal*, 336(7658), 1413–1415.
- Jacobsen, H., Moe, V., Ivarsson, T., Wentzel-Larsen, T., & Smith, L. (2013). Cognitive development and social-emotional functioning in young foster children: A follow-up study from 2 to 3 years of age. *Child Psychiatry and Human Development*, 44(5), 666–677. <https://doi.org/10.1007/s10578-013-0360-3>.
- James, S., Landsverk, J., & Slymen, D. J. (2004). Placement movement in out-of-home care: Patterns and predictors. *Children and Youth Services Review*, 26(2), 185–206. <https://doi.org/10.1016/j.chilcyouth.2004.01.008>.
- Jones, R., Everson-Hock, E. S., Papaioannou, D., Guillaume, L., Goyder, E., Chilcott, J., et al. (2011). Factors associated with outcomes for looked-after children and young people: A correlates review of the literature. *Child: Care, Health and Development*, 37(5), 613–622.
- Kelly, J. F., & Korfmacher, J. (2008). *Raising a baby*. Unpublished manuscript, University of Washington.
- Kendall, J. M. (2003). Designing a research project: Randomised controlled trials and their principles. *Emergency Medicine Journal*, 20(2), 164–168. <https://doi.org/10.1136/emj.20.2.164>.
- Kerr, L., & Cossar, J. (2014). Attachment interventions with foster and adoptive parents: A systematic review. *Child Abuse Review*, 23(6), 426–439. <https://doi.org/10.1002/car.2313>.
- Kim, H. K., & Leve, L. D. (2011). Substance use and delinquency among middle school girls in foster care: A three-year follow-up of a randomized controlled trial (Randomized Controlled Trial; Research Support, N.I.H., Extramural). *Journal of Consulting and Clinical Psychology*, 79(6), 740–750. <https://doi.org/10.1037/a0025949>.
- Kim, H. K., Pears, K. C., Leve, L. D., Chamberlain, P., & Smith, D. K. (2013). Intervention effects on health-risking sexual behavior among girls in foster care: The role of placement disruption and tobacco and marijuana use. *Journal of Child & Adolescent Substance Abuse*, 22(5), 370–387. <https://doi.org/10.1080/1067828X.2013.788880>.
- Kinsey, D., & Schlösser, A. (2013). Interventions in foster and kinship care: A systematic review. *Clinical Child Psychology and Psychiatry*, 18(3), 429–463.
- Koren-Karie, N., & Markman-Gefen, R. (2016). Foster caregiver insightfulness and emotional investment in foster children. *Journal of Social Work*, 16(4), 489–509. <https://doi.org/10.1177/1468017315581528>.
- Kramer, L. (2010). Quality of sibling interaction coding manual: More fun with sisters and brothers 2007–2010. University of Illinois at Urbana-Champaign, Urbana, IL. Unpublished manuscript.
- Kramer, L., & Houston, D. (1999). Hope for the children: A community-based approach to supporting families who adopt children with special needs. *Child Welfare: Journal of Policy, Practice, and Program*, 78(5), 611–635.
- Leenarts, L. E. W., Diehle, J., Doreleijers, T. A. H., Jansma, E. P., & Lindauer, R. J. L. (2013). Evidence-based treatments for children with trauma-related psychopathology as a result of childhood maltreatment: A systematic review. *European Child and Adolescent Psychiatry*, 22(5), 269–283. <https://doi.org/10.1007/s00787-012-0367-5>.
- Leve, L. D., Harold, G. T., Chamberlain, P., Landsverk, J. A., Fisher, P. A., & Vostanis, P. (2012). Practitioner review: Children in foster care—Vulnerabilities and evidence-based interventions that promote resilience processes. *Journal of Child Psychology and Psychiatry*, 53(12), 1197–1211. <https://doi.org/10.1111/j.1469-7610.2012.02594.x>.
- Lewis-Morrarty, E., Dozier, M., Bernard, K., Terracciano, S. M., & Moore, S. V. (2012). Cognitive flexibility and theory of mind

- outcomes among foster children: preschool follow-up results of a randomized clinical trial (Randomized Controlled Trial; Research Support, N.I.H., Extramural; Research Support, Non-U.S. Gov't). *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, 51(2), S17–S22. <https://doi.org/10.1016/j.jadohealth.2012.05.005>.
- Linares, L. (2008). *The sibling aggression scale*. Unpublished scale. New York, NY: NYU Child Study Center, New York University.
- Linares, L. O., Montalto, D., Li, M., & Oza, V. S. (2006). A promising parenting intervention in foster care. *Journal of Consulting and Clinical Psychology*, 74(1), 32–41. <https://doi.org/10.1037/0022-006X.74.1.32>.
- Locke, L. M., & Prinz, R. J. (2002). Measurement of parental discipline and nurturance. *Clinical Psychology Review*, 22(6), 895–929. [https://doi.org/10.1016/S0272-7358\(02\)00133-2](https://doi.org/10.1016/S0272-7358(02)00133-2).
- Luke, N., Sinclair, I., Woolgar, M., & Sebba, J. (2014). *What works in preventing and treating poor mental health in looked after children*. London: NSPCC.
- Maaskant, A. M., van Rooij, F. B., Overbeek, G. J., Oort, F. J., & Hermanns, J. M. A. (2016). Parent training in foster families with children with behavior problems: Follow-up results from a randomized controlled trial. *Children and Youth Services Review*. <https://doi.org/10.1016/j.childyouth.2016.09.005>.
- Macdonald, G. M., & Turner, W. (2005). An experiment in helping foster-carers manage challenging behaviour. *British Journal of Social Work*, 35(8), 1265–1282. <https://doi.org/10.1093/bjsw/bch204>.
- Marti Haidar, Y. M. (2013). *What is the experience of foster care mothers?* ProQuest Dissertations Publishing.
- McWayne, C. M., Owsianik, M., Green, L. E., & Fantuzzo, J. W. (2008). Parenting behaviors and preschool children's social and emotional skills: A question of the consequential validity of traditional parenting constructs for low-income African Americans. *Early Childhood Research Quarterly*, 23(2), 173–192.
- Mersky, J. P., Topitzes, J., Grant-Savelle, S. D., Brondino, M. J., & McNeil, C. B. (2016). Adapting parent–child interaction therapy to foster care: Outcomes from a randomized trial. *Research on Social Work Practice*, 26(2), 157.
- Mersky, J. P., Topitzes, J., Janczewski, C. E., & McNeil, C. B. (2015). Enhancing foster parent training with parent–child interaction therapy: Evidence from a randomized field experiment. *Journal of the Society for Social Work and Research*, 6(4), 591–616. <https://doi.org/10.1086/684123>.
- Millward, R., Kennedy, E., Towlson, K., & Minnis, H. (2006). Reactive attachment disorder in looked-after children. *Emotional and Behavioural Difficulties*, 11(4), 273–279.
- Milner, J. (1990). *An interpretive manual for the child abuse potential inventory*. DeKalb, IL: Psytec. Inc.
- Minnis, H., & Devine, C. (2001). The effect of foster carer training on the emotional and behavioural functioning of looked after children. *Adoption and Fostering*, 25(1), 44–54.
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), 1.
- Nissen, S. B., Magidson, T., Gross, K., & Bergstrom, C. T. (2016). Publication bias and the canonization of false facts. *ELife*. <https://doi.org/10.7554/eLife.21451>.
- O'Dell, S. L., Tarler-Benlolo, L., & Flynn, J. M. (1979). An instrument to measure knowledge of behavioral principles as applied to children. *Journal of Behavior Therapy and Experimental Psychiatry*, 10(1), 29–34.
- Oosterman, M., Schuengel, C., Wim Slot, N., Bullens, R. A. R., & Doreleijers, T. A. H. (2007). Disruptions in foster care: A review and meta-analysis. *Children and Youth Services Review*, 29(1), 53–76. <https://doi.org/10.1016/j.childyouth.2006.07.003>.
- Opris, D., & Macavei, B. (2007). The profile of emotional distress; Norms for the Romanian population. *Journal of Cognitive & Behavioral Psychotherapies*, 7(2), 139.
- Oswald, D. P., & Ollendick, T. H. (1989). Role taking and social competence in autism and mental retardation. *Journal of Autism and Developmental Disorders*, 19(1), 119–127. <https://doi.org/10.1007/bf02212723>.
- Pears, K. C., Healey, C. V., Fisher, P. A., Braun, D., Gill, C., Conte, H. M., et al. (2014). Immediate effects of a program to promote school readiness in low-income children: Results of a pilot study. *Education and Treatment of Children*, 37(3), 431–460.
- Pears, K. C., Kim, H. K., & Fisher, P. A. (2012). Effects of a school readiness intervention for children in foster care on oppositional and aggressive behaviors in kindergarten. *Children and Youth Services Review*, 34(12), 2361–2366. <https://doi.org/10.1016/j.childyouth.2012.08.015>.
- Pears, K. C., Kim, H. K., & Fisher, P. A. (2016). Decreasing risk factors for later alcohol use and antisocial behaviors in children in foster care by increasing early promotive factors. *Children and Youth Services Review*, 65, 156–165. <https://doi.org/10.1016/j.childyouth.2016.04.005>.
- Pears, K. C., Kim, H. K., Fisher, P. A., & Yoerger, K. (2013). Early school engagement and late elementary outcomes for maltreated children in foster care. *Developmental Psychology*, 49(12), 2201–2211. <https://doi.org/10.1037/a0032218>.
- Perry, B. D. (2009). Examining child maltreatment through a neurodevelopmental lens: Clinical applications of the neurosequential model of therapeutics. *Journal of Loss and Trauma*, 14(4), 240–255. <https://doi.org/10.1080/15325020903004350>.
- Price, J. M., Chamberlain, P., Landsverk, J., Reid, J. B., Leve, L. D., & Laurent, H. (2008). Effects of a foster parent training intervention on placement changes of children in foster care. *Child Maltreatment*, 13(1), 64–75. <https://doi.org/10.1177/1077559507310612>.
- Price, J. M., Roesch, S., Walsh, N. E., & Landsverk, J. (2015). Effects of the KEEP foster parent intervention on child and sibling behavior problems and parental stress during a randomized implementation trial. *Prevention Science*, 16(5), 685–695. <https://doi.org/10.1007/s11121-014-0532-9>.
- Racusin, R., Maerlender, A. C., Sengupta, A., Isquith, P. K., & Straus, M. B. (2005). Psychosocial treatment of children in foster care: A review. *Community Mental Health Journal*, 41(2), 199–221. <https://doi.org/10.1007/s10597-005-2656-7>.
- Redding, R. E., Fried, C., & Britner, P. A. (2000). Predictors of placement outcomes in treatment foster care: Implications for foster parent selection and service delivery. *Journal of Child and Family Studies*, 9(4), 425–447. <https://doi.org/10.1023/a:1009418809133>.
- Rothwell, P. M. (2005). External validity of randomised controlled trials: “To whom do the results of this trial apply?”. *The Lancet*, 365(9453), 82–93. [https://doi.org/10.1016/S0140-6736\(04\)17670-8](https://doi.org/10.1016/S0140-6736(04)17670-8).
- Roweton, W. E. (1990). Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). Explaining delinquency and drug use. Beverly Hills, CA: Sage, 176 pp., \$25.00. *Psychology in the Schools*, 27(1), 91–92. [https://doi.org/10.1002/1520-6807\(199001\)27:1<91:AID-PITS2310270115>3.0.CO;2-Y](https://doi.org/10.1002/1520-6807(199001)27:1<91:AID-PITS2310270115>3.0.CO;2-Y).
- Rubin, D. M., O'Reilly, A. L. R., Luan, X. Q., & Localio, A. R. (2007). The impact of placement stability on behavioral well-being for children in foster care. *Pediatrics*, 119(2), 336–344. <https://doi.org/10.1542/peds.2006-1995>.
- Ryan, J. P., & Testa, M. F. (2005). Child maltreatment and juvenile delinquency: Investigating the role of placement and placement instability. *Children and Youth Services Review*, 27(3), 227–249. <https://doi.org/10.1016/j.childyouth.2004.05.007>.
- Sackett, D. L., Rosenberg, W. M. C., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What



- it is and what it isn't. *BMJ*, 312(7023), 71–72. <https://doi.org/10.1136/bmj.312.7023.71>.
- Schuengel, C., Oosterman, M., & Sterkenburg, P. S. (2009). Children with disrupted attachment histories: Interventions and psychophysiological indices of effects. *Child and Adolescent Psychiatry and Mental Health*. <https://doi.org/10.1186/1753-2000-3-26>.
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, 143(2), 534.
- Sinclair, I., Baker, C., Wilson, K., & Gibbs, I. (2005). *Foster Children: Where they go and how they get on*. London: Jessica Kingsley Publishers.
- Smith, D. K., Leve, L. D., & Chamberlain, P. (2011). Preventing internalizing and externalizing problems in girls in foster care as they enter middle school: Impact of an intervention. *Prevention Science*, 12(3), 269–277. <https://doi.org/10.1007/s1121-011-0211-z>.
- Smith, J., & Ross, H. (2007). Training parents to mediate sibling disputes affects children's negotiation and conflict understanding. *Child Development*, 78(3), 790–805.
- Spieker, S., Nelson, E. M., & Condon, M.-C. (2011). Validity of the TAS-45 as a measure of toddler–parent attachment: preliminary evidence from Early Head Start families. *Attachment & Human Development*, 13(1), 69–90.
- Spieker, S. J., Oxford, M. L., & Fleming, C. B. (2014). Permanency outcomes for toddlers in child welfare 2 years after a randomized trial of a parenting intervention. *Children and Youth Services Review*, 44, 201–206. <https://doi.org/10.1016/j.childyouth.2014.06.017>.
- Spieker, S. J., Oxford, M. L., Kelly, J. F., Nelson, E. M., & Fleming, C. B. (2012). Promoting first relationships: Randomized trial of a relationship-based intervention for toddlers in child welfare. *Child Maltreatment*, 17(4), 271–286. <https://doi.org/10.1177/1077559512458176>.
- Sprang, G. (2009). The efficacy of a relational treatment for maltreated children and their families. *Child and Adolescent Mental Health*, 14(2), 81–88. <https://doi.org/10.1111/j.1475-3588.2008.00499.x>.
- Tarren-Sweeney, M. (2013a). An investigation of complex attachment- and trauma-related symptomatology among children in foster and kinship care. *Child Psychiatry and Human Development*, 44(6), 727–741. <https://doi.org/10.1007/s10578-013-0366-x>.
- Tarren-Sweeney, M. (2013b). The assessment checklist for adolescents—ACA: A scale for measuring the mental health of young people in foster, kinship, residential and adoptive care. *Children and Youth Services Review*, 35(3), 384–393. <https://doi.org/10.1016/j.childyouth.2012.12.005>.
- Tarren-Sweeney, M. (2014). Our twenty-first century quest: Locating effective mental health interventions for children and young people in care, and those adopted from care. In M. Tarren-Sweeney & A. Vetere (Eds.), *Mental health services for vulnerable children and young people: Supporting children who are, or have been, in foster care*. New York: Routledge.
- Turner, W., Macdonald, G. M., & Dennis, J. A. (2007). Behavioural and cognitive behavioural training interventions for assisting foster carers in the management of difficult behaviour. *Cochrane Database of Systematic Reviews*. <https://doi.org/10.1002/14651858.CD003760.pub3>.
- Van Andel, H., Post, W., Jansen, L., Van der Gaag, R. J., Knorth, E., & Grietens, H. (2016). Optimizing foster family placement for infants and toddlers: A randomized controlled trial on the effect of the foster family intervention. *American Journal of Orthopsychiatry*, 86(3), 332–344. <https://doi.org/10.1037/ort0000162>.
- Webster-Stratton, C. (1998). Preventing conduct problems in head start children: Strengthening parenting competencies. *Journal of Consulting and Clinical Psychology*, 66(5), 715–730. <https://doi.org/10.1037/0022-006X.66.5.715>.
- West, S. L., Gartlehner, G., Mansfield, A. J., Poole, C., Tant, E., Lenfestey, N., et al. (2010). *Comparative effectiveness review methods: Clinical heterogeneity*. <https://www.ncbi.nlm.nih.gov/books/NBK53310/>.
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1(1), 297.